

EDA del volumen de alimentos consumidos por los hogares españoles

David Moreno

2023-09-10

Introducción

En el marco del proyecto de análisis de datos, hemos explorado una extensa colección de información que abarca dos décadas, desde el año 2000 hasta el 2022, relacionada con los hábitos de consumo de alimentos en los hogares españoles.

Este conjunto de datos se encuentra organizado en 22 archivos Excel, cada uno correspondiente a un año específico, y está disponible en un repositorio de GitHub. Cada archivo contiene múltiples hojas de datos, y una de ellas se centra en el “volumen”, que mide la cantidad de alimentos consumidos por los hogares españoles.

En este contexto, emprendemos un Análisis Exploratorio de Datos (EDA) con el objetivo de comprender más profundamente este conjunto de datos y extraer insights significativos.

1. Preparación del entorno de trabajo

Configuro el repositorio de CRAN de manera no interactiva

```
options(repos = c(CRAN = "https://cran.r-project.org"))
```

Instalo las librerías que voy a utilizar

```
install.packages("readxl")
```

```
## Installing package into 'C:/Users/gilga/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)
## package 'readxl' successfully unpacked and MD5 sums checked
## Warning: cannot remove prior installation of package 'readxl'
## Warning in file.copy(savedcopy, lib, recursive = TRUE): problema al copiar
## C:\Users\gilga\AppData\Local\R\win-library\4.3\00LOCK\readxl\libs\x64\readxl.dll
## a C:\Users\gilga\AppData\Local\R\win-library\4.3\readxl\libs\x64\readxl.dll:
## Permission denied
## Warning: restored 'readxl'
##
## The downloaded binary packages are in
## C:\Users\gilga\AppData\Local\Temp\Rtmp2bYRhB\downloaded_packages
```

```
install.packages("httr")
```

```
## Installing package into 'C:/Users/gilga/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)
## package 'httr' successfully unpacked and MD5 sums checked
##
```

```
## The downloaded binary packages are in
## C:\Users\gilga\AppData\Local\Temp\Rtmp2bYRhB\downloaded_packages
install.packages("tidyverse")

## Installing package into 'C:/Users/gilga/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)
## package 'tidyverse' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
## C:\Users\gilga\AppData\Local\Temp\Rtmp2bYRhB\downloaded_packages
install.packages("rmarkdown")

## Installing package into 'C:/Users/gilga/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)
## package 'rmarkdown' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
## C:\Users\gilga\AppData\Local\Temp\Rtmp2bYRhB\downloaded_packages
install.packages("dplyr")

## Installing package into 'C:/Users/gilga/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)
## package 'dplyr' successfully unpacked and MD5 sums checked
## Warning: cannot remove prior installation of package 'dplyr'
## Warning in file.copy(savedcopy, lib, recursive = TRUE): problema al copiar
## C:\Users\gilga\AppData\Local\R\win-library\4.3\00LOCK\dplyr\libs\x64\dplyr.dll
## a C:\Users\gilga\AppData\Local\R\win-library\4.3\dplyr\libs\x64\dplyr.dll:
## Permission denied
## Warning: restored 'dplyr'
##
## The downloaded binary packages are in
## C:\Users\gilga\AppData\Local\Temp\Rtmp2bYRhB\downloaded_packages
install.packages("corrplot")

## Installing package into 'C:/Users/gilga/AppData/Local/R/win-library/4.3'
## (as 'lib' is unspecified)
## package 'corrplot' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
## C:\Users\gilga\AppData\Local\Temp\Rtmp2bYRhB\downloaded_packages
```

Y, por último, se cargan las librerías.

```
library(readxl)
library(httr)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
```

```
## v ggplot2 3.4.3      v tibble 3.2.1
## v lubridate 1.9.2    v tidyr 1.3.0
## v purrr 1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag() masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(rmarkdown)
library(dplyr)
library(corrplot)
```

```
## corrplot 0.92 loaded
```

2. Obtención de los datos

2.1 Extracción de los datos

```
# Creo una lista vacía donde almacenaré los df.
lista_df_volumen <- list()

# Mediante un bucle for, recorro cada archivo, para extraer los datos
for (i in seq(from = 1, to = 23, by = 1)) {

  # Datos de la ruta
  ruta <- "https://github.com/DMorgon/portafolios/raw/main"
  rama <- "alimentacion/datos_origen"

  # Creo la ruta desde donde se descargaran los archivos
  archivo_url <- paste0(ruta, "/", rama, "/", i + 1999, ".xlsx")

  # Creo el nombre del archivo local
  archivo_local <- paste0(i, ".xlsx")

  # Descargar el archivo desde la URL
  response <- GET(archivo_url, write_disk(archivo_local, overwrite = TRUE))

  # Creo el df con los datos del archivo local
  df <- read_excel(archivo_local, sheet = 3, skip = 2)

  #Agrego el df a la lista de df.
  lista_df_volumen[[i]] <- df

  # Borro el archivo local
  file.remove(archivo_local)
}
```

```
## New names:
## New names:
## New names:
## New names:
## New names:
## New names:
## New names:
## New names:
```

```
## New names:
## New names:
## New names:
## New names:
## New names:
## New names:
## New names:
## * `` -> `...1`

## Warning in file.remove(archivo_local): no fue posible abrir el archivo
## '15.xlsx', motivo 'Permission denied'

## New names:
## * `` -> `...1`

## Warning in file.remove(archivo_local): no fue posible abrir el archivo
## '16.xlsx', motivo 'Permission denied'

## New names:
## New names:
## New names:
## New names:
## New names:
## New names:
## * `` -> `...1`

## Warning in file.remove(archivo_local): no fue posible abrir el archivo
## '22.xlsx', motivo 'Permission denied'

## New names:
## * `` -> `...1`

## Warning in file.remove(archivo_local): no fue posible abrir el archivo
## '23.xlsx', motivo 'Permission denied'

# Elimino las variables que ya no se utilizará.
rm(ruta, rama, archivo_url, archivo_local, response, df, i)
```

2.2 Exploración inicial de los datos

```
# Mediante un bucle for, imprimo la estructura inter de cada dataframe
for (i in seq(1, 23)) {

  # Imprimo el nombre de cada dataframe
  cat("Nombre de la tabla de datos: df_volumen_", i+1999, "\n")

  #Imprimo la estructura interna de cada dataframe
  cat(str(lista_df_volumen[[i]]), "\n")
}

## Nombre de la tabla de datos: df_volumen_ 2000
## tibble [334 x 19] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:334] "TOTAL ALIMENTACION" "HUEVOS KGS" "HUEVOS" "GALLINA" ...
## $ .TOTAL ESPAÑA : num [1:334] 25489859 422362 6646199 6590734 55464 ...
## $ CATALUÑA : num [1:334] 4112919 62709 992946 977394 15552 ...
## $ ARAGON : num [1:334] 792586 16595 261159 258957 2202 ...
## $ BALEARES : num [1:334] 439345 5991 93771 93573 0 ...
## $ VALENCIA : num [1:334] 2567574 40248 630606 628545 2060 ...
```

```

## $ MURCIA : num [1:334] 741976 11614 181802 181415 0 ...
## $ ANDALUCIA : num [1:334] 4563527 73883 1163013 1152831 10182 ...
## $ MADRID : num [1:334] 2999485 48692 766816 759709 7106 ...
## $ CASTILLA-LA MANCHA: num [1:334] 1166605 20035 314941 312695 2247 ...
## $ EXTREMADURA : num [1:334] 727689 12638 198541 197264 0 ...
## $ CASTILLA Y LEON : num [1:334] 1805147 37559 589167 586438 2729 ...
## $ GALICIA : num [1:334] 1696696 20725 324393 323724 670 ...
## $ ASTURIAS : num [1:334] 810800 17018 266632 265768 0 ...
## $ CANTABRIA : num [1:334] 296146 6057 95143 94541 0 ...
## $ PAIS VASCO : num [1:334] 1290681 26584 421938 414165 7773 ...
## $ RIOJA : num [1:334] 165249 3344 52615 52182 0 ...
## $ NAVARRA : num [1:334] 338831 6753 106435 105341 0 ...
## $ CANARIAS : num [1:334] 974603 11917 186280 186193 0 ...
##
## Nombre de la tabla de datos: df_volumen_ 2001
## tibble [334 x 19] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:334] "TOTAL ALIMENTACION" "HUEVOS KGS" "HUEVOS" "GALLINA" ...
## $ .TOTAL ESPAÑA : num [1:334] 25908528 417681 6572606 6517694 54912 ...
## $ CATALUÑA : num [1:334] 4096287 59887 947018 933650 13368 ...
## $ ARAGON : num [1:334] 821838 18658 293080 291245 1835 ...
## $ BALEARES : num [1:334] 466386 6444 101135 100605 531 ...
## $ VALENCIA : num [1:334] 2640091 37239 584483 581369 3114 ...
## $ MURCIA : num [1:334] 724401 10724 167998 167489 509 ...
## $ ANDALUCIA : num [1:334] 4731316 74063 1168728 1155109 13619 ...
## $ MADRID : num [1:334] 2951362 46723 734684 729194 5489 ...
## $ CASTILLA-LA MANCHA: num [1:334] 1166196 19656 307635 307024 611 ...
## $ EXTREMADURA : num [1:334] 718590 13592 212906 212274 0 ...
## $ CASTILLA Y LEON : num [1:334] 1877070 37815 593812 590313 0 ...
## $ GALICIA : num [1:334] 1806662 24567 385201 383604 1597 ...
## $ ASTURIAS : num [1:334] 786473 16923 264906 264328 578 ...
## $ CANTABRIA : num [1:334] 290138 5883 92100 91888 212 ...
## $ PAIS VASCO : num [1:334] 1293727 24054 381550 374791 6759 ...
## $ RIOJA : num [1:334] 164051 3416 54356 53199 0 ...
## $ NAVARRA : num [1:334] 356810 8066 126980 125856 1124 ...
## $ CANARIAS : num [1:334] 1017129 9971 156035 155757 0 ...
##
## Nombre de la tabla de datos: df_volumen_ 2002
## tibble [361 x 19] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:361] "TOTAL ALIMENTACION" "HUEVOS KGS" "HUEVOS" "GALLINA" ...
## $ .TOTAL ESPAÑA : num [1:361] 26464420 416069 6553333 6491406 61927 ...
## $ CATALUÑA : num [1:361] 4214577 58395 922880 910478 12401 ...
## $ ARAGON : num [1:361] 895550 18136 286319 282832 3488 ...
## $ BALEARES : num [1:361] 483387 7524 118029 117482 547 ...
## $ VALENCIA : num [1:361] 2651358 36911 580454 576046 4408 ...
## $ MURCIA : num [1:361] 746333 8922 140076 139290 787 ...
## $ ANDALUCIA : num [1:361] 4906711 72196 1142737 1125343 17395 ...
## $ MADRID : num [1:361] 3060406 46288 730010 721990 8020 ...
## $ CASTILLA-LA MANCHA: num [1:361] 1164925 19678 308627 307258 1369 ...
## $ EXTREMADURA : num [1:361] 669252 13122 205423 204964 460 ...
## $ CASTILLA Y LEON : num [1:361] 1851659 37103 582045 579304 2741 ...
## $ GALICIA : num [1:361] 1861194 28416 445344 443759 0 ...
## $ ASTURIAS : num [1:361] 802864 15394 241084 240429 654 ...
## $ CANTABRIA : num [1:361] 293221 5135 80285 80225 0 ...
## $ PAIS VASCO : num [1:361] 1291499 25557 402780 398697 4082 ...

```

```

## $ RIOJA : num [1:361] 183215 3826 61070 59551 1519 ...
## $ NAVARRA : num [1:361] 335237 7343 115457 114596 861 ...
## $ CANARIAS : num [1:361] 1053032 12122 190712 189162 0 ...
##
## Nombre de la tabla de datos: df_volumen_ 2003
## tibble [361 x 19] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:361] "TOTAL ALIMENTACION" "HUEVOS KGS" "HUEVOS" "GALLINA" ...
## $ .TOTAL ESPAÑA : num [1:361] 26947770 425611 6707958 6639476 68482 ...
## $ CATALUÑA : num [1:361] 4405903 59883 951080 932821 18258 ...
## $ ARAGON : num [1:361] 938700 16946 267525 264278 3247 ...
## $ BALEARES : num [1:361] 524271 7353 115647 114743 904 ...
## $ VALENCIA : num [1:361] 2698249 39108 615144 610307 4836 ...
## $ MURCIA : num [1:361] 762893 8088 126961 126267 694 ...
## $ ANDALUCIA : num [1:361] 4983794 76343 1206609 1190309 16300 ...
## $ MADRID : num [1:361] 3147504 48238 760384 752488 7896 ...
## $ CASTILLA-LA MANCHA: num [1:361] 1176489 20705 324862 323269 1593 ...
## $ EXTREMADURA : num [1:361] 633660 13386 209186 209157 0 ...
## $ CASTILLA Y LEON : num [1:361] 1872239 37665 591239 588008 3231 ...
## $ GALICIA : num [1:361] 1790479 27546 433621 429815 3806 ...
## $ ASTURIAS : num [1:361] 826143 15645 244712 244406 306 ...
## $ CANTABRIA : num [1:361] 321059 5738 89780 89635 0 ...
## $ PAIS VASCO : num [1:361] 1308029 24847 391197 387685 3511 ...
## $ RIOJA : num [1:361] 185644 3849 60968 59982 986 ...
## $ NAVARRA : num [1:361] 311529 7241 113383 113097 0 ...
## $ CANARIAS : num [1:361] 1061183 13030 205661 203208 2453 ...
##
## Nombre de la tabla de datos: df_volumen_ 2004
## tibble [448 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:448] ".TOTAL ALIMENTACION" "HUEVOS KGS" "HUEVOS" "GALLINA" ...
## $ T.ESPAÑA : num [1:448] 27842842 414273 6534733 6461592 73142 ...
## $ CATALUÑA : num [1:448] 4473621 56638 899501 882282 17218 ...
## $ ARAGON : num [1:448] 909011 14539 234627 225794 8833 ...
## $ BALEARES : num [1:448] 554057 6067 95214 94723 0 ...
## $ VALENCIA : num [1:448] 2901364 39836 625709 621825 3884 ...
## $ MURCIA : num [1:448] 790034 9265 145272 144673 599 ...
## $ T.ANDALUCIA : num [1:448] 5306345 77613 1228573 1209771 18802 ...
## $ MADRID : num [1:448] 3231757 47605 748851 742904 5947 ...
## $ CASTILLA LA MANCHA: num [1:448] 1230857 20211 317914 315399 2515 ...
## $ EXTREMADURA : num [1:448] 696837 12729 199266 198824 0 ...
## $ CASTILLA LEON : num [1:448] 1811700 32069 503495 500639 2856 ...
## $ GALICIA : num [1:448] 1806492 28990 453964 452789 1175 ...
## $ ASTURIAS : num [1:448] 757310 13597 213393 212272 0 ...
## $ CANTABRIA : num [1:448] 359595 7188 112307 112307 0 ...
## $ PAIS VASCO : num [1:448] 1349676 22853 362439 356077 6363 ...
## $ LA RIOJA : num [1:448] 176564 3757 59519 58560 959 ...
## $ NAVARRA : num [1:448] 355810 7404 116200 115596 604 ...
## $ CANARIAS : num [1:448] 1131812 13912 218493 217160 1333 ...
## $ NORESTE : num [1:448] 5936689 77245 1229342 1202799 26543 ...
## $ LEVANTE : num [1:448] 3691398 49101 770980 766498 4483 ...
## $ ANDALUCIA : num [1:448] 5306345 77613 1228573 1209771 18802 ...
## $ CENTRO-SUR : num [1:448] 5159451 80545 1266031 1257126 8904 ...
## $ CASTILLA Y LEON : num [1:448] 1811700 32069 503495 500639 2856 ...
## $ NOROESTE : num [1:448] 2563802 42587 667357 665062 2295 ...
## $ NORTE : num [1:448] 2241645 41202 650466 642540 7926 ...

```

```

## $ T.CANARIAS      : num [1:448] 1131812 13912 218493 217160 1333 ...
##
## Nombre de la tabla de datos: df_volumen_ 2005
## tibble [465 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1            : chr [1:465] ".TOTAL ALIMENTACION" "HUEVOS KGS" "HUEVOS" "GALLINA" ...
## $ T.ESPAÑA        : num [1:465] 28035987 405986 6415650 6330177 85473 ...
## $ CATALUÑA        : num [1:465] 4473097 58136 928190 904709 23481 ...
## $ ARAGON          : num [1:465] 877249 14054 227304 218162 9143 ...
## $ BALEARES        : num [1:465] 579408 6667 104799 104062 737 ...
## $ VALENCIA        : num [1:465] 2892051 37696 594613 587964 6650 ...
## $ MURCIA          : num [1:465] 712751 8984 141037 140250 788 ...
## $ T.ANDALUCIA     : num [1:465] 5267875 71924 1137357 1121296 16061 ...
## $ MADRID          : num [1:465] 3224823 46649 734615 727822 6793 ...
## $ CASTILLA LA MANCHA: num [1:465] 1284187 19201 302630 299523 3107 ...
## $ EXTREMADURA    : num [1:465] 712209 12278 192197 191771 426 ...
## $ CASTILLA LEON   : num [1:465] 1831716 30794 485045 480436 4609 ...
## $ GALICIA         : num [1:465] 1931543 29001 454229 452948 1281 ...
## $ ASTURIAS        : num [1:465] 749559 12860 201463 200831 632 ...
## $ CANTABRIA       : num [1:465] 399644 6864 107592 107180 411 ...
## $ PAIS VASCO      : num [1:465] 1354861 24505 388858 381782 7076 ...
## $ LA RIOJA        : num [1:465] 185481 3669 58252 57150 1102 ...
## $ NAVARRA         : num [1:465] 395691 8138 127954 127014 939 ...
## $ CANARIAS        : num [1:465] 1163826 14568 229515 227277 2238 ...
## $ NORESTE         : num [1:465] 5929753 78857 1260294 1226933 33361 ...
## $ LEVANTE         : num [1:465] 3604802 46680 735651 728214 7437 ...
## $ ANDALUCIA       : num [1:465] 5267875 71924 1137357 1121296 16061 ...
## $ CENTRO-SUR      : num [1:465] 5221219 78127 1229441 1219115 10326 ...
## $ CASTILLA Y LEON : num [1:465] 1831716 30794 485045 480436 4609 ...
## $ NOROESTE        : num [1:465] 2681103 41861 655692 653779 1913 ...
## $ NORTE           : num [1:465] 2335677 43175 682655 673127 9528 ...
## $ T.CANARIAS      : num [1:465] 1163826 14568 229515 227277 2238 ...
##
## Nombre de la tabla de datos: df_volumen_ 2006
## tibble [465 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1            : chr [1:465] ".TOTAL ALIMENTACION" "HUEVOS KGS" "HUEVOS" "GALLINA" ...
## $ T.ESPAÑA        : num [1:465] 28171562 396400 6268547 6179895 88652 ...
## $ CATALUÑA        : num [1:465] 4530833 56648 904191 881590 22600 ...
## $ ARAGON          : num [1:465] 885494 14029 227142 217742 9401 ...
## $ BALEARES        : num [1:465] 559067 6038 95792 94068 1724 ...
## $ VALENCIA        : num [1:465] 2869997 37343 590208 582241 7967 ...
## $ MURCIA          : num [1:465] 730272 9279 146494 144713 1781 ...
## $ T.ANDALUCIA     : num [1:465] 5408332 73157 1157475 1140413 17062 ...
## $ MADRID          : num [1:465] 3317417 46539 733177 726067 7110 ...
## $ CASTILLA LA MANCHA: num [1:465] 1283247 18629 295752 290213 5539 ...
## $ EXTREMADURA    : num [1:465] 713222 11827 185254 184714 540 ...
## $ CASTILLA LEON   : num [1:465] 1904119 30079 472482 469523 2958 ...
## $ GALICIA         : num [1:465] 1952117 29109 455321 454729 592 ...
## $ ASTURIAS        : num [1:465] 778049 13289 207903 207591 311 ...
## $ CANTABRIA       : num [1:465] 423810 6446 101884 100498 1386 ...
## $ PAIS VASCO      : num [1:465] 1357996 22886 362762 356644 6118 ...
## $ LA RIOJA        : num [1:465] 192006 3463 55074 53938 1136 ...
## $ NAVARRA         : num [1:465] 398717 7556 118879 117904 975 ...
## $ CANARIAS        : num [1:465] 866869 10082 158758 157309 1449 ...
## $ NORESTE        : num [1:465] 5975395 76715 1227125 1193400 33725 ...

```

```

## $ LEVANTE : num [1:465] 3600268 46623 736702 726955 9748 ...
## $ ANDALUCIA : num [1:465] 5408332 73157 1157475 1140413 17062 ...
## $ CENTRO-SUR : num [1:465] 5313886 76995 1214183 1200994 13190 ...
## $ CASTILLA Y LEON : num [1:465] 1904119 30079 472482 469523 2958 ...
## $ NOROESTE : num [1:465] 2730166 42398 663224 662320 904 ...
## $ NORTE : num [1:465] 2372530 40351 638599 628983 9616 ...
## $ T.CANARIAS : num [1:465] 866869 10082 158758 157309 1449 ...
##
## Nombre de la tabla de datos: df_volumen_ 2007
## tibble [552 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:552] ".TOTAL ALIMENTACION" "HUEVOS KGS" "HUEVOS" "HUEVOS G." ...
## $ T.ESPAÑA : num [1:552] 28896142 396047 6266617 6173718 559294 ...
## $ CATALUÑA : num [1:552] 4589318 55812 892470 868274 45282 ...
## $ ARAGON : num [1:552] 900872 12998 209368 201936 18776 ...
## $ BALEARES : num [1:552] 549529 5980 94517 93241 8627 ...
## $ VALENCIA : num [1:552] 2946060 38943 614574 607364 22755 ...
## $ MURCIA : num [1:552] 816098 10216 162305 159120 13140 ...
## $ T.ANDALUCIA : num [1:552] 5421590 70753 1121294 1102602 129170 ...
## $ MADRID : num [1:552] 3326341 46007 725277 717677 20695 ...
## $ CASTILLA LA MANCHA : num [1:552] 1326982 19746 311109 308062 48121 ...
## $ EXTREMADURA : num [1:552] 727958 11573 181703 180659 32539 ...
## $ CASTILLA LEON : num [1:552] 1869969 29512 464986 460414 54357 ...
## $ GALICIA : num [1:552] 2015140 27743 434340 433322 85937 ...
## $ ASTURIAS : num [1:552] 764099 13233 207266 206675 29179 ...
## $ CANTABRIA : num [1:552] 446999 6231 99239 97004 3573 ...
## $ PAIS VASCO : num [1:552] 1365144 22773 361217 354838 25712 ...
## $ LA RIOJA : num [1:552] 193123 3591 57143 55915 1453 ...
## $ NAVARRA : num [1:552] 377799 6600 104335 102900 2609 ...
## $ CANARIAS : num [1:552] 1259121 14335 225476 223714 17370 ...
## $ NORESTE : num [1:552] 6039719 74790 1196354 1163451 72685 ...
## $ LEVANTE : num [1:552] 3762158 49159 776880 766484 35895 ...
## $ ANDALUCIA : num [1:552] 5421590 70753 1121294 1102602 129170 ...
## $ CENTRO-SUR : num [1:552] 5381281 77326 1218089 1206399 101355 ...
## $ CASTILLA Y LEON : num [1:552] 1869969 29512 464986 460414 54357 ...
## $ NOROESTE : num [1:552] 2779239 40976 641605 639997 115116 ...
## $ NORTE : num [1:552] 2383064 39195 621935 610657 33347 ...
## $ T.CANARIAS : num [1:552] 1259121 14335 225476 223714 17370 ...
##
## Nombre de la tabla de datos: df_volumen_ 2008
## tibble [556 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:556] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL HUEVOS" ...
## $ T.ESPAÑA : num [1:556] 30299348 388250 6173632 6046555 752637 ...
## $ CATALUÑA : num [1:556] 4937489 56155 902505 872771 50114 ...
## $ ARAGON : num [1:556] 931002 12274 200959 190085 23721 ...
## $ BALEARES : num [1:556] 686274 6554 104471 102027 17366 ...
## $ VALENCIA : num [1:556] 3233859 39776 630631 619804 39065 ...
## $ MURCIA : num [1:556] 903028 9619 152633 149860 10931 ...
## $ T.ANDALUCIA : num [1:556] 5185089 64005 1019506 996477 143140 ...
## $ MADRID : num [1:556] 3695784 47877 760095 745851 57909 ...
## $ CASTILLA LA MANCHA : num [1:556] 1427990 19539 308287 304734 75244 ...
## $ EXTREMADURA : num [1:556] 702744 11228 177656 175035 59843 ...
## $ CASTILLA LEON : num [1:556] 1887795 28508 452076 444212 56549 ...
## $ GALICIA : num [1:556] 1941849 24693 386913 385625 109034 ...
## $ ASTURIAS : num [1:556] 798925 11930 187199 186254 34049 ...

```



```

## $ CANTABRIA      : num [1:556] 416093 6917 110346 107665 8481 ...
## $ PAIS VASCO     : num [1:556] 1449360 23697 377394 368945 34642 ...
## $ LA RIOJA       : num [1:556] 204541 3312 53095 51505 1595 ...
## $ NAVARRA        : num [1:556] 375356 5765 90789 89946 3177 ...
## $ CANARIAS       : num [1:556] 1522169 16402 259076 255760 27778 ...
## $ NORESTE        : num [1:556] 6554765 74983 1207936 1164883 91201 ...
## $ LEVANTE        : num [1:556] 4136886 49394 783264 769664 49996 ...
## $ ANDALUCIA      : num [1:556] 5185089 64005 1019506 996477 143140 ...
## $ CENTRO-SUR     : num [1:556] 5826519 78644 1246038 1225619 192995 ...
## $ CASTILLA Y LEON : num [1:556] 1887795 28508 452076 444212 56549 ...
## $ NOROESTE       : num [1:556] 2740774 36623 574112 571879 143083 ...
## $ NORTE          : num [1:556] 2445351 39692 631624 618061 47894 ...
## $ T.CANARIAS     : num [1:556] 1522169 16402 259076 255760 27778 ...
##
## Nombre de la tabla de datos: df_volumen_ 2009
## tibble [556 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1           : chr [1:556] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL HUEVOS UNDS."
## $ T.ESPAÑA       : num [1:556] 30798560 403505 6415812 6284206 840545 ...
## $ CATALUÑA       : num [1:556] 5009432 58674 940914 912304 68817 ...
## $ ARAGON         : num [1:556] 942122 12410 201956 192421 32033 ...
## $ BALEARES       : num [1:556] 715874 8217 130692 127971 28592 ...
## $ VALENCIA       : num [1:556] 3339073 41884 666025 652291 31452 ...
## $ MURCIA         : num [1:556] 919299 10781 171176 167956 16538 ...
## $ T.ANDALUCIA    : num [1:556] 5287891 69871 1112806 1087826 199697 ...
## $ MADRID         : num [1:556] 3779685 49456 784006 770667 54750 ...
## $ CASTILLA LA MANCHA: num [1:556] 1462772 19741 312759 307658 57290 ...
## $ EXTREMADURA   : num [1:556] 681051 10019 158344 156207 30995 ...
## $ CASTILLA LEON  : num [1:556] 1928983 29955 474912 466775 95894 ...
## $ GALICIA        : num [1:556] 1963944 26216 411589 409256 113372 ...
## $ ASTURIAS       : num [1:556] 785406 10812 169650 168812 25288 ...
## $ CANTABRIA      : num [1:556] 406436 7087 111906 110511 14668 ...
## $ PAIS VASCO     : num [1:556] 1417557 22652 361684 352507 35202 ...
## $ LA RIOJA       : num [1:556] 212917 3383 54291 52586 1130 ...
## $ NAVARRA        : num [1:556] 383405 6215 99268 96718 2739 ...
## $ CANARIAS       : num [1:556] 1562713 16132 253833 251739 32088 ...
## $ NORESTE        : num [1:556] 6667428 79301 1273562 1232695 129442 ...
## $ LEVANTE        : num [1:556] 4258372 52665 837201 820247 47990 ...
## $ ANDALUCIA      : num [1:556] 5287891 69871 1112806 1087826 199697 ...
## $ CENTRO-SUR     : num [1:556] 5923508 79216 1255109 1234532 143035 ...
## $ CASTILLA Y LEON : num [1:556] 1928983 29955 474912 466775 95894 ...
## $ NOROESTE       : num [1:556] 2749350 37028 581239 578069 138660 ...
## $ NORTE          : num [1:556] 2420316 39337 627149 612322 53739 ...
## $ T.CANARIAS     : num [1:556] 1562713 16132 253833 251739 32088 ...
##
## Nombre de la tabla de datos: df_volumen_ 2010
## tibble [556 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1           : chr [1:556] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL HUEVOS UNDS."
## $ T.ESPAÑA       : num [1:556] 30448957 378445 6025170 5892475 712586 ...
## $ CATALUÑA       : num [1:556] 4974058 55981 899535 870099 57066 ...
## $ ARAGON         : num [1:556] 918410 11555 188235 179120 20138 ...
## $ BALEARES       : num [1:556] 717320 8235 130640 128311 15777 ...
## $ VALENCIA       : num [1:556] 3417402 39958 635995 622188 31181 ...
## $ MURCIA         : num [1:556] 890614 10247 162422 159675 11263 ...
## $ T.ANDALUCIA    : num [1:556] 5357829 64974 1041305 1010387 154321 ...

```

```

## $ MADRID : num [1:556] 3844423 49922 789870 778204 65628 ...
## $ CASTILLA LA MANCHA: num [1:556] 1399451 19077 301610 297430 64310 ...
## $ EXTREMADURA : num [1:556] 659162 8564 135009 133598 27983 ...
## $ CASTILLA LEON : num [1:556] 1844445 26888 426940 418859 79385 ...
## $ GALICIA : num [1:556] 1886547 23019 361335 359368 81445 ...
## $ ASTURIAS : num [1:556] 716718 10051 157984 156866 25049 ...
## $ CANTABRIA : num [1:556] 389098 6632 104197 103523 8618 ...
## $ PAIS VASCO : num [1:556] 1396776 20539 328273 319555 31619 ...
## $ LA RIOJA : num [1:556] 208637 2865 46814 44378 2163 ...
## $ NAVARRA : num [1:556] 369579 5617 89190 87510 3072 ...
## $ CANARIAS : num [1:556] 1458488 14322 225815 223404 33567 ...
## $ NORESTE : num [1:556] 6609788 75771 1218411 1177531 92981 ...
## $ LEVANTE : num [1:556] 4308016 50205 798417 781864 42444 ...
## $ ANDALUCIA : num [1:556] 5357829 64974 1041305 1010387 154321 ...
## $ CENTRO-SUR : num [1:556] 5903037 77563 1226488 1209231 157921 ...
## $ CASTILLA Y LEON : num [1:556] 1844445 26888 426940 418859 79385 ...
## $ NOROESTE : num [1:556] 2603266 33070 519319 516234 106493 ...
## $ NORTE : num [1:556] 2364090 35653 568475 554966 45473 ...
## $ T.CANARIAS : num [1:556] 1458488 14322 225815 223404 33567 ...
##
## Nombre de la tabla de datos: df_volumen_ 2011
## tibble [556 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:556] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL HUEVOS UNDS."
## $ T.ESPAÑA : num [1:556] 30238963 376893 6010732 5866395 664712 ...
## $ CATALUÑA : num [1:556] 4902582 54236 879281 841548 45205 ...
## $ ARAGON : num [1:556] 909834 11150 181915 172784 18146 ...
## $ BALEARES : num [1:556] 737923 8393 133044 130782 9936 ...
## $ VALENCIA : num [1:556] 3378902 41252 656570 642336 30571 ...
## $ MURCIA : num [1:556] 918952 10547 169102 163993 9018 ...
## $ T.ANDALUCIA : num [1:556] 5322351 64526 1030578 1004080 132209 ...
## $ MADRID : num [1:556] 3837011 49519 783613 771911 55861 ...
## $ CASTILLA LA MANCHA: num [1:556] 1386758 19131 301996 298351 59285 ...
## $ EXTREMADURA : num [1:556] 685281 9044 143499 140902 36227 ...
## $ CASTILLA LEON : num [1:556] 1825840 27455 434530 427952 89384 ...
## $ GALICIA : num [1:556] 1839357 21770 341544 339906 84237 ...
## $ ASTURIAS : num [1:556] 721604 9801 153926 152999 19704 ...
## $ CANTABRIA : num [1:556] 388470 6355 100103 99156 9381 ...
## $ PAIS VASCO : num [1:556] 1356384 20314 327035 315615 34742 ...
## $ LA RIOJA : num [1:556] 206551 2921 46962 45392 1815 ...
## $ NAVARRA : num [1:556] 364881 5644 92649 87363 3857 ...
## $ CANARIAS : num [1:556] 1456292 14835 234385 231325 25135 ...
## $ NORESTE : num [1:556] 6550339 73779 1194240 1145114 73287 ...
## $ LEVANTE : num [1:556] 4297854 51798 825672 806329 39589 ...
## $ ANDALUCIA : num [1:556] 5322351 64526 1030578 1004080 132209 ...
## $ CENTRO-SUR : num [1:556] 5909050 77694 1229108 1211164 151373 ...
## $ CASTILLA Y LEON : num [1:556] 1825840 27455 434530 427952 89384 ...
## $ NOROESTE : num [1:556] 2560961 31572 495470 492904 103941 ...
## $ NORTE : num [1:556] 2316286 35234 566749 547527 49794 ...
## $ T.CANARIAS : num [1:556] 1456292 14835 234385 231325 25135 ...
##
## Nombre de la tabla de datos: df_volumen_ 2012
## tibble [556 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:556] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL HUEVOS UNDS."
## $ T.ESPAÑA : num [1:556] 30436538 380675 6085999 5922501 619533 ...

```

```
## $ CATALUÑA      : num [1:556] 4887099 55014 890065 853947 53834 ...
## $ ARAGON        : num [1:556] 903054 11283 184880 174715 18623 ...
## $ BALEARES      : num [1:556] 755783 8538 135720 132969 10581 ...
## $ VALENCIA      : num [1:556] 3341924 43065 685116 670628 38401 ...
## $ MURCIA        : num [1:556] 1014472 12088 193199 188066 11139 ...
## $ T.ANDALUCIA   : num [1:556] 5381179 66340 1064139 1031460 121957 ...
## $ MADRID        : num [1:556] 3812896 49285 784444 767420 46813 ...
## $ CASTILLA LA MANCHA: num [1:556] 1383580 18360 291661 285983 45678 ...
## $ EXTREMADURA  : num [1:556] 677791 7837 123924 122177 17596 ...
## $ CASTILLA LEON : num [1:556] 1857711 26110 414466 406774 63974 ...
## $ GALICIA       : num [1:556] 1845789 21485 339172 335068 94580 ...
## $ ASTURIAS      : num [1:556] 730165 9717 152577 151683 20589 ...
## $ CANTABRIA     : num [1:556] 405236 6627 105145 103244 10070 ...
## $ PAIS VASCO    : num [1:556] 1365679 20503 330146 318556 37263 ...
## $ LA RIOJA      : num [1:556] 199362 2684 43644 41630 2561 ...
## $ NAVARRA       : num [1:556] 392517 6257 102514 96884 5241 ...
## $ CANARIAS      : num [1:556] 1482307 15482 245189 241297 20634 ...
## $ NORESTE       : num [1:556] 6545936 74835 1210664 1161632 83037 ...
## $ LEVANTE       : num [1:556] 4356396 55153 878314 858694 49539 ...
## $ ANDALUCIA     : num [1:556] 5381179 66340 1064139 1031460 121957 ...
## $ CENTRO-SUR    : num [1:556] 5874266 75482 1200029 1175579 110087 ...
## $ CASTILLA Y LEON : num [1:556] 1857711 26110 414466 406774 63974 ...
## $ NOROESTE      : num [1:556] 2575954 31202 491749 486751 115169 ...
## $ NORTE         : num [1:556] 2362793 36071 581449 560314 55135 ...
## $ T.CANARIAS    : num [1:556] 1482307 15482 245189 241297 20634 ...
```

```
##
```

```
## Nombre de la tabla de datos: df_volumen_ 2013
```

```
## tibble [580 x 27] (S3: tbl_df/tbl/data.frame)
```

```
## $ ...1          : chr [1:580] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA      : num [1:580] 30359818 390414 6225347 6077050 5393308 ...
## $ CATALUÑA      : num [1:580] 4983530 57721 926795 897287 826193 ...
## $ ARAGÓN        : num [1:580] 922865 12083 198439 187017 168402 ...
## $ ILLES BALEARS : num [1:580] 724332 8744 139635 136069 125073 ...
## $ COMUNITAT VALENCIANA : num [1:580] 3322293 43904 698300 683727 651486 ...
## $ REGIÓN DE MURCIA : num [1:580] 918727 10565 168609 164420 148640 ...
## $ ANDALUCÍA     : num [1:580] 5383984 67352 1077243 1047771 906797 ...
## $ COMUNIDAD DE MADRID : num [1:580] 3788530 50629 804945 788503 745924 ...
## $ CASTILLA - LA MANCHA : num [1:580] 1404196 18871 300512 293813 253656 ...
## $ EXTREMADURA  : num [1:580] 691324 8357 131681 130381 106796 ...
## $ CASTILLA Y LEÓN : num [1:580] 1789955 25520 404834 397632 341673 ...
## $ GALICIA       : num [1:580] 1849191 22301 350523 348070 240502 ...
## $ PRINCIPADO DE ASTURIAS: num [1:580] 752630 10882 171556 169752 143284 ...
## $ CANTABRIA     : num [1:580] 399844 7344 116217 114471 99667 ...
## $ PAIS VASCO    : num [1:580] 1347104 21006 335404 326891 283420 ...
## $ LA RIOJA      : num [1:580] 207928 2973 47966 46180 40689 ...
## $ C. FORAL DE NAVARRA : num [1:580] 421312 7185 116086 111551 103205 ...
## $ CANARIAS      : num [1:580] 1452074 14976 236605 233516 207900 ...
## $ NORESTE       : num [1:580] 6630726 78549 1264869 1220372 1119668 ...
## $ LEVANTE       : num [1:580] 4241020 54469 866909 848147 800127 ...
## $ ANDALUCIA     : num [1:580] 5383984 67352 1077243 1047771 906797 ...
## $ CENTRO-SUR    : num [1:580] 5884050 77857 1237137 1212697 1106377 ...
## $ CASTILLA Y LEON : num [1:580] 1789955 25520 404834 397632 341673 ...
## $ NOROESTE      : num [1:580] 2601821 33183 522078 517822 383786 ...
## $ NORTE         : num [1:580] 2376188 38508 615672 599093 526982 ...
```

```

## $ T.CANARIAS : num [1:580] 1452074 14976 236605 233516 207900 ...
##
## Nombre de la tabla de datos: df_volumen_ 2014
## tibble [580 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:580] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:580] 29639336 379524 6051722 5907524 5196661 ...
## $ CATALUÑA : num [1:580] 4722704 56422 905371 877183 800395 ...
## $ ARAGÓN : num [1:580] 886422 11914 194416 184626 156915 ...
## $ ILLES BALEARS : num [1:580] 764076 9160 145522 142678 129099 ...
## $ COMUNITAT VALENCIANA : num [1:580] 3287455 43515 691876 677714 636366 ...
## $ REGIÓN DE MURCIA : num [1:580] 901808 10381 165278 161640 139335 ...
## $ ANDALUCÍA : num [1:580] 5188320 64097 1025286 997122 845128 ...
## $ COMUNIDAD DE MADRID : num [1:580] 3699811 48570 770208 756805 713244 ...
## $ CASTILLA - LA MANCHA : num [1:580] 1362260 18037 289816 280353 238072 ...
## $ EXTREMADURA : num [1:580] 686768 8117 128877 126447 102539 ...
## $ CASTILLA Y LEÓN : num [1:580] 1739367 23572 375231 367038 315352 ...
## $ GALICIA : num [1:580] 1844454 23148 364475 361175 252556 ...
## $ PRINCIPADO DE ASTURIAS: num [1:580] 752088 10467 165451 163186 141067 ...
## $ CANTABRIA : num [1:580] 391073 6738 107068 104947 94810 ...
## $ PAIS VASCO : num [1:580] 1352056 20664 330603 321443 280775 ...
## $ LA RIOJA : num [1:580] 201601 2974 47403 46298 42201 ...
## $ C. FORAL DE NAVARRA : num [1:580] 391890 6452 103010 100404 91680 ...
## $ CANARIAS : num [1:580] 1467185 15295 241831 238467 217128 ...
## $ NORESTE : num [1:580] 6373201 77495 1245309 1204487 1086409 ...
## $ LEVANTE : num [1:580] 4189263 53897 857153 839354 775701 ...
## $ ANDALUCIA : num [1:580] 5188320 64097 1025286 997122 845128 ...
## $ CENTRO-SUR : num [1:580] 5748838 74724 1188901 1163604 1053856 ...
## $ CASTILLA Y LEON : num [1:580] 1739367 23572 375231 367038 315352 ...
## $ NOROESTE : num [1:580] 2596542 33615 529926 524361 393623 ...
## $ NORTE : num [1:580] 2336620 36828 588085 573092 509466 ...
## $ T.CANARIAS : num [1:580] 1467185 15295 241831 238467 217128 ...
##
## Nombre de la tabla de datos: df_volumen_ 2015
## tibble [580 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:580] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:580] 29249368 376634 6012237 5861319 5102986 ...
## $ CATALUÑA : num [1:580] 4790762 57201 917409 889379 816124 ...
## $ ARAGÓN : num [1:580] 874257 12456 205650 192584 158868 ...
## $ ILLES BALEARS : num [1:580] 728887 8442 134103 131495 116863 ...
## $ COMUNITAT VALENCIANA : num [1:580] 3172025 43380 691777 675235 625483 ...
## $ REGIÓN DE MURCIA : num [1:580] 891713 10923 174499 169971 147434 ...
## $ ANDALUCÍA : num [1:580] 5185824 65065 1040285 1012269 842813 ...
## $ COMUNIDAD DE MADRID : num [1:580] 3629497 49068 778256 764554 716955 ...
## $ CASTILLA - LA MANCHA : num [1:580] 1378805 17617 282196 273991 233627 ...
## $ EXTREMADURA : num [1:580] 695257 8375 132332 130592 105358 ...
## $ CASTILLA Y LEÓN : num [1:580] 1698756 22682 364206 352597 301753 ...
## $ GALICIA : num [1:580] 1805696 22713 358241 354269 220700 ...
## $ PRINCIPADO DE ASTURIAS: num [1:580] 686620 9547 150378 148945 123116 ...
## $ CANTABRIA : num [1:580] 357437 5953 93635 92907 82659 ...
## $ PAIS VASCO : num [1:580] 1316096 19167 308825 297756 271458 ...
## $ LA RIOJA : num [1:580] 183975 2475 39827 38451 35518 ...
## $ C. FORAL DE NAVARRA : num [1:580] 378994 5737 91345 89320 82160 ...
## $ CANARIAS : num [1:580] 1474768 15831 249274 247003 222097 ...
## $ NORESTE : num [1:580] 6393907 78098 1257161 1213458 1091856 ...

```

```

## $ LEVANTE : num [1:580] 4063738 54304 866276 845206 772917 ...
## $ ANDALUCIA : num [1:580] 5185824 65065 1040285 1012269 842813 ...
## $ CENTRO-SUR : num [1:580] 5703559 75061 1192784 1169137 1055940 ...
## $ CASTILLA Y LEON : num [1:580] 1698756 22682 364206 352597 301753 ...
## $ NOROESTE : num [1:580] 2492316 32260 508619 503214 343816 ...
## $ NORTE : num [1:580] 2236501 33332 533631 518434 471794 ...
## $ T.CANARIAS : num [1:580] 1474768 15831 249274 247003 222097 ...
##
## Nombre de la tabla de datos: df_volumen_ 2016
## tibble [580 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:580] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:580] 29035111 376172 5995698 5855836 5078030 ...
## $ CATALUÑA : num [1:580] 4783230 58256 932882 906051 812372 ...
## $ ARAGÓN : num [1:580] 884085 13349 219141 206619 170737 ...
## $ ILLES BALEARs : num [1:580] 731036 8658 137707 134831 121778 ...
## $ COMUNITAT VALENCIANA : num [1:580] 3081883 41252 657345 642197 597709 ...
## $ REGIÓN DE MURCIA : num [1:580] 898855 10597 167580 165201 149696 ...
## $ ANDALUCÍA : num [1:580] 5139244 64546 1034035 1003800 832186 ...
## $ COMUNIDAD DE MADRID : num [1:580] 3564012 48617 768801 757935 719182 ...
## $ CASTILLA - LA MANCHA : num [1:580] 1343559 17055 272990 265282 228784 ...
## $ EXTREMADURA : num [1:580] 669839 8216 130506 127984 105957 ...
## $ CASTILLA Y LEÓN : num [1:580] 1650505 23467 373222 365454 312058 ...
## $ GALICIA : num [1:580] 1819634 21616 341296 337092 198380 ...
## $ PRINCIPADO DE ASTURIAS: num [1:580] 741396 10444 165167 162816 136031 ...
## $ CANTABRIA : num [1:580] 343294 5911 93440 92168 77223 ...
## $ PAIS VASCO : num [1:580] 1319972 20072 320446 312370 278198 ...
## $ LA RIOJA : num [1:580] 170781 2656 42538 41308 38071 ...
## $ C. FORAL DE NAVARRA : num [1:580] 388774 5624 88936 87683 79821 ...
## $ CANARIAS : num [1:580] 1505012 15837 249664 247045 219847 ...
## $ NORESTE : num [1:580] 6398351 80262 1289731 1247501 1104886 ...
## $ LEVANTE : num [1:580] 3980738 51849 824926 807397 747406 ...
## $ ANDALUCIA : num [1:580] 5139244 64546 1034035 1003800 832186 ...
## $ CENTRO-SUR : num [1:580] 5577410 73888 1172297 1151201 1053924 ...
## $ CASTILLA Y LEON : num [1:580] 1650505 23467 373222 365454 312058 ...
## $ NOROESTE : num [1:580] 2561030 32060 506463 499908 334411 ...
## $ NORTE : num [1:580] 2222821 34264 545361 533529 473313 ...
## $ T.CANARIAS : num [1:580] 1505012 15837 249664 247045 219847 ...
##
## Nombre de la tabla de datos: df_volumen_ 2017
## tibble [580 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:580] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:580] 28833331 382752 6104662 5957507 5241213 ...
## $ CATALUÑA : num [1:580] 4747910 58926 943992 916412 836163 ...
## $ ARAGÓN : num [1:580] 854413 12270 201704 189861 163957 ...
## $ ILLES BALEARs : num [1:580] 745888 8855 141906 137703 123624 ...
## $ COMUNITAT VALENCIANA : num [1:580] 3092157 42492 674525 661983 629463 ...
## $ REGIÓN DE MURCIA : num [1:580] 929124 11564 183769 180120 162909 ...
## $ ANDALUCÍA : num [1:580] 5156465 65675 1053451 1021124 860817 ...
## $ COMUNIDAD DE MADRID : num [1:580] 3518970 48705 773479 758709 732710 ...
## $ CASTILLA - LA MANCHA : num [1:580] 1333060 16941 271378 263472 238521 ...
## $ EXTREMADURA : num [1:580] 648255 8432 133406 131438 108682 ...
## $ CASTILLA Y LEÓN : num [1:580] 1632477 23559 374768 366876 316419 ...
## $ GALICIA : num [1:580] 1820363 24402 385796 380447 221831 ...
## $ PRINCIPADO DE ASTURIAS: num [1:580] 699113 9837 154924 153484 132511 ...

```

```

## $ CANTABRIA : num [1:580] 346198 5965 93910 93070 78813 ...
## $ PAIS VASCO : num [1:580] 1294324 20621 331236 320526 292229 ...
## $ LA RIOJA : num [1:580] 163213 2584 41749 40114 35664 ...
## $ C. FORAL DE NAVARRA : num [1:580] 383002 5954 93889 92874 84881 ...
## $ CANARIAS : num [1:580] 1468402 15970 250780 249293 222019 ...
## $ NORESTE : num [1:580] 6348210 80051 1287601 1243976 1123744 ...
## $ LEVANTE : num [1:580] 4021280 54057 858294 842103 792372 ...
## $ ANDALUCIA : num [1:580] 5156465 65675 1053451 1021124 860817 ...
## $ CENTRO-SUR : num [1:580] 5500285 74078 1178263 1153620 1079913 ...
## $ CASTILLA Y LEON : num [1:580] 1632477 23559 374768 366876 316419 ...
## $ NOROESTE : num [1:580] 2519475 34240 540721 533932 354342 ...
## $ NORTE : num [1:580] 2186737 35123 560783 546583 491587 ...
## $ T.CANARIAS : num [1:580] 1468402 15970 250780 249293 222019 ...
##
## Nombre de la tabla de datos: df_volumen_ 2018
## tibble [580 x 27] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:580] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:580] 28775131 385508 6151575 5999862 5322049 ...
## $ CATALUÑA : num [1:580] 4768749 61876 990854 962362 881773 ...
## $ ARAGÓN : num [1:580] 833377 12180 198104 188864 157721 ...
## $ ILLES BALEARs : num [1:580] 751088 9106 145520 141677 129364 ...
## $ COMUNITAT VALENCIANA : num [1:580] 3100131 42242 670887 658020 623857 ...
## $ REGIÓN DE MURCIA : num [1:580] 916319 12310 194658 191909 175215 ...
## $ ANDALUCÍA : num [1:580] 5173502 68040 1092006 1057785 895656 ...
## $ COMUNIDAD DE MADRID : num [1:580] 3543859 49212 785240 765914 740366 ...
## $ CASTILLA - LA MANCHA : num [1:580] 1329160 16079 255332 250482 229634 ...
## $ EXTREMADURA : num [1:580] 662988 8850 139864 137980 111675 ...
## $ CASTILLA Y LEÓN : num [1:580] 1555899 22563 359759 351210 308542 ...
## $ GALICIA : num [1:580] 1837383 23593 375369 367390 227876 ...
## $ PRINCIPADO DE ASTURIAS : num [1:580] 683028 9344 147472 145731 131066 ...
## $ CANTABRIA : num [1:580] 343979 5818 91375 90828 76230 ...
## $ PAIS VASCO : num [1:580] 1310177 20450 328935 317791 299172 ...
## $ LA RIOJA : num [1:580] 166744 2362 38263 36663 32154 ...
## $ C. FORAL DE NAVARRA : num [1:580] 340097 4726 74454 73722 69118 ...
## $ CANARIAS : num [1:580] 1458652 16758 263483 261534 232628 ...
## $ NORESTE : num [1:580] 6353214 83162 1334479 1292903 1168859 ...
## $ LEVANTE : num [1:580] 4016450 54552 865545 849929 799072 ...
## $ ANDALUCIA : num [1:580] 5173502 68040 1092006 1057785 895656 ...
## $ CENTRO-SUR : num [1:580] 5536007 74141 1180436 1154376 1081676 ...
## $ CASTILLA Y LEON : num [1:580] 1555899 22563 359759 351210 308542 ...
## $ NOROESTE : num [1:580] 2520411 32937 522841 513120 358943 ...
## $ NORTE : num [1:580] 2160997 33357 533027 519005 476674 ...
## $ T.CANARIAS : num [1:580] 1458652 16758 263483 261534 232628 ...
##
## Nombre de la tabla de datos: df_volumen_ 2019
## tibble [686 x 19] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:686] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:686] 28669381 384218 6131955 5979603 5322619 ...
## $ CATALUÑA : num [1:686] 4851289 61685 988529 959261 899416 ...
## $ ARAGÓN : num [1:686] 803460 11759 191699 182261 157283 ...
## $ ILLES BALEARs : num [1:686] 793440 8822 141682 137140 128065 ...
## $ COMUNITAT VALENCIANA : num [1:686] 3095376 41726 664093 649728 609546 ...
## $ REGIÓN DE MURCIA : num [1:686] 870154 11522 182831 179510 169002 ...
## $ ANDALUCÍA : num [1:686] 5061991 67260 1078765 1045778 891858 ...

```

```

## $ COMUNIDAD DE MADRID : num [1:686] 3579605 50246 800961 782162 746377 ...
## $ CASTILLA - LA MANCHA : num [1:686] 1281526 15298 243795 238145 218043 ...
## $ EXTREMADURA : num [1:686] 667395 8346 131773 130157 104850 ...
## $ CASTILLA Y LEÓN : num [1:686] 1549569 22070 352816 343358 295042 ...
## $ GALICIA : num [1:686] 1775887 24332 385596 379177 233320 ...
## $ PRINCIPADO DE ASTURIAS: num [1:686] 666467 8873 139454 138492 121440 ...
## $ CANTABRIA : num [1:686] 350084 5775 90766 90140 79153 ...
## $ PAIS VASCO : num [1:686] 1293909 21450 342720 333764 316120 ...
## $ LA RIOJA : num [1:686] 174199 2653 42851 41196 38393 ...
## $ C. FORAL DE NAVARRA : num [1:686] 362719 5441 86107 84820 78636 ...
## $ CANARIAS : num [1:686] 1492309 16959 267517 264514 236075 ...
##
## Nombre de la tabla de datos: df_volumen_ 2020
## tibble [681 x 19] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:681] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:681] 31878711 449770 7170438 7001219 169219 ...
## $ CATALUÑA : num [1:681] 5337307 74282 1186350 1155905 30445 ...
## $ ARAGÓN : num [1:681] 915169 14586 236898 226237 10661 ...
## $ ILLES BALEARs : num [1:681] 875008 10989 176305 170857 5449 ...
## $ COMUNITAT VALENCIANA : num [1:681] 3405733 47294 752895 736388 16507 ...
## $ REGIÓN DE MURCIA : num [1:681] 980193 13246 210782 206255 4528 ...
## $ ANDALUCÍA : num [1:681] 5655873 77126 1234709 1199602 35107 ...
## $ COMUNIDAD DE MADRID : num [1:681] 4064034 61603 981961 958959 23002 ...
## $ CASTILLA - LA MANCHA : num [1:681] 1407353 17677 281036 275312 5724 ...
## $ EXTREMADURA : num [1:681] 721467 8999 142890 140195 2695 ...
## $ CASTILLA Y LEÓN : num [1:681] 1687600 23764 377555 370163 7392 ...
## $ GALICIA : num [1:681] 1916380 27822 439628 433806 5821 ...
## $ PRINCIPADO DE ASTURIAS: num [1:681] 764300 11309 177808 176498 1310 ...
## $ CANTABRIA : num [1:681] 369830 6053 95121 94481 641 ...
## $ PAIS VASCO : num [1:681] 1471163 25420 408488 395090 13398 ...
## $ LA RIOJA : num [1:681] 198713 3179 50608 49505 1103 ...
## $ C. FORAL DE NAVARRA : num [1:681] 416563 6991 110224 109046 1177 ...
## $ CANARIAS : num [1:681] 1692026 19430 307180 302921 4259 ...
##
## Nombre de la tabla de datos: df_volumen_ 2021
## tibble [681 x 19] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:681] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:681] 29586612 404499 6459668 6294484 165184 ...
## $ CATALUÑA : num [1:681] 4880281 63972 1024046 995022 29023 ...
## $ ARAGÓN : num [1:681] 845522 12785 208839 198089 10750 ...
## $ ILLES BALEARs : num [1:681] 827354 9432 152075 146510 5565 ...
## $ COMUNITAT VALENCIANA : num [1:681] 3217745 44019 702248 685122 17126 ...
## $ REGIÓN DE MURCIA : num [1:681] 928159 11816 188314 183944 4370 ...
## $ ANDALUCÍA : num [1:681] 5226166 68547 1098842 1065893 32949 ...
## $ COMUNIDAD DE MADRID : num [1:681] 3699939 55424 880713 863270 17443 ...
## $ CASTILLA - LA MANCHA : num [1:681] 1332921 17573 279435 273682 5752 ...
## $ EXTREMADURA : num [1:681] 671405 8047 127867 125344 2523 ...
## $ CASTILLA Y LEÓN : num [1:681] 1578683 21620 347312 336056 11256 ...
## $ GALICIA : num [1:681] 1784974 24264 384664 378107 6556 ...
## $ PRINCIPADO DE ASTURIAS: num [1:681] 713180 9891 155147 154435 712 ...
## $ CANTABRIA : num [1:681] 344863 5683 89506 88661 846 ...
## $ PAIS VASCO : num [1:681] 1392179 23675 382656 367559 15096 ...
## $ LA RIOJA : num [1:681] 180236 2621 41932 40781 1152 ...
## $ C. FORAL DE NAVARRA : num [1:681] 381589 6346 100554 98893 1662 ...

```

```
## $ CANARIAS : num [1:681] 1581415 18783 295518 293116 2402 ...
##
## Nombre de la tabla de datos: df_volumen_ 2022
## tibble [689 x 19] (S3: tbl_df/tbl/data.frame)
## $ ...1 : chr [1:689] ".TOTAL ALIMENTACION" "T.HUEVOS KGS" "T.HUEVOS UNDS." "TOTAL I
## $ T.ESPAÑA : num [1:689] 26987664 375190 5988041 5839062 148979 ...
## $ CATALUÑA : num [1:689] 4488271 58921 946353 915879 30474 ...
## $ ARAGÓN : num [1:689] 733188 11100 179983 172226 7757 ...
## $ ILLES BALEARS : num [1:689] 714256 7078 114175 109925 4250 ...
## $ COMUNITAT VALENCIANA : num [1:689] 2965964 42210 671905 657236 14669 ...
## $ REGIÓN DE MURCIA : num [1:689] 870634 10618 169616 165214 4402 ...
## $ ANDALUCÍA : num [1:689] 4797066 65692 1050341 1022018 28322 ...
## $ COMUNIDAD DE MADRID : num [1:689] 3405536 50918 807585 793369 14216 ...
## $ CASTILLA - LA MANCHA : num [1:689] 1232720 15834 253902 246209 7693 ...
## $ EXTREMADURA : num [1:689] 617207 7019 111356 109362 1994 ...
## $ CASTILLA Y LEÓN : num [1:689] 1509767 20525 327295 319490 7805 ...
## $ GALICIA : num [1:689] 1626094 21805 347290 339479 7811 ...
## $ PRINCIPADO DE ASTURIAS: num [1:689] 643616 9299 146224 145125 1099 ...
## $ CANTABRIA : num [1:689] 308103 5743 90220 89645 576 ...
## $ PAIS VASCO : num [1:689] 1257759 23049 371029 358122 12907 ...
## $ LA RIOJA : num [1:689] 164739 2825 44925 43992 933 ...
## $ C. FORAL DE NAVARRA : num [1:689] 343987 6001 94795 93571 1224 ...
## $ CANARIAS : num [1:689] 1308757 16553 261048 258201 2846 ...
##
# Elimino la variable i que no se utilizará
rm(i)
```

Después de llevar a cabo una exploración inicial de la estructura interna de cada dataframe, se han identificado las siguientes conclusiones:

- Se han encontrado un total de 22 dataframes.
- Existe una disparidad en la cantidad de registros, es decir, en la variedad de alimentos, entre los distintos dataframes.
- Algunos dataframes contienen 19 variables, mientras que otros presentan 27. Esto se debe a que, en los dataframes correspondientes a los años de 2004 a 2018, se incorporaron 8 regiones adicionales.
- Se ha observado variabilidad en los nombres de las variables en algunos dataframes.
- También es posible que los nombres de los alimentos varíen de un dataframe a otro.

En consecuencia, las tareas que se llevarán a cabo a continuación son las siguientes:

- **Normalización de las variables:** Dado que existen diferencias en el número de variables y en sus nombres, se procederá a eliminar las 8 variables relacionadas con las regiones y a corregir el nombre de las variables restantes para asegurar consistencia.
- **Creación de un conjunto de datos consolidado:** Una vez normalizados las variables, se combinarán todos los dataframes en un conjunto consolidado, especificando además el año al que corresponde cada registro.

2.3 Transformación de los datos

Elimino las variables sobrantes

```
for (i in seq(5, 19)) {
  if (i <= 13) {
    lista_df_volumen[[i]] <- lista_df_volumen[[i]][, !colnames(lista_df_volumen[[i]]) %in% c("T.ANDALUCIA", "T.CANTABRIA", "T.CASTILLA Y LEON", "T.CASTILLA - LA MANCHA", "T.COMUNITAT VALENCIANA", "T.COMUNIDAD DE MADRID", "T.EXTREMADURA", "T.GALICIA", "T.ILLES BALEARS", "T.PRINCIPADO DE ASTURIAS", "T.PAIS VASCO", "T.REGIÓN DE MURCIA", "T.TOTAL ALIMENTACION", "T.TOTAL HUEVOS KGS", "T.TOTAL HUEVOS UNDS")]
  }
}
```



```
df_volumen <- df_volumen %>%
  pivot_longer(cols = -c(PRODUCTOS, AÑO),
               names_to = "REGIONES",
               values_to = "VOLUMEN")
```

2.4 Carga de los datos

Hago una copia de df_volumen llamada data, con la que trabajaré a continuación

```
data <- df_volumen
```

3. Analisis descriptivo

En primer lugar, obtengo una vista descriptiva sobre el contenido del conjunto de datos,

```
print(head(data))
```

```
## # A tibble: 6 x 4
##   AÑO PRODUCTOS REGIONES VOLUMEN
##   <dbl> <chr>      <chr>      <dbl>
## 1  2000 TOTAL ALIMENTACION ANDALUCIA 4563527.
## 2  2000 TOTAL ALIMENTACION ARAGON    792586.
## 3  2000 TOTAL ALIMENTACION ASTURIAS  810800.
## 4  2000 TOTAL ALIMENTACION BALEARES  439345.
## 5  2000 TOTAL ALIMENTACION CANARIAS  974603.
## 6  2000 TOTAL ALIMENTACION CANTABRIA 296146.
```

Examino la estructura interna del conjunto de los datos.

```
str(data)
```

```
## tibble [221,706 x 4] (S3: tbl_df/tbl/data.frame)
## $ AÑO      : num [1:221706] 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 2000 ...
## $ PRODUCTOS: chr [1:221706] "TOTAL ALIMENTACION" "TOTAL ALIMENTACION" "TOTAL ALIMENTACION" "TOTAL ALIMENTACION" ...
## $ REGIONES : chr [1:221706] "ANDALUCIA" "ARAGON" "ASTURIAS" "BALEARES" ...
## $ VOLUMEN  : num [1:221706] 4563527 792586 810800 439345 974603 ...
```

Ahora, obtengo un resumen estadístico de cada una de las variables.

```
summary(data)
```

```
##      AÑO      PRODUCTOS      REGIONES      VOLUMEN
## Min.   :2000   Length:221706   Length:221706   Min.    :    0
## 1st Qu.:2007   Class :character   Class :character 1st Qu.:   275
## Median :2013   Mode  :character   Mode  :character Median :  1451
## Mean   :2012                                     Mean  :  28538
## 3rd Qu.:2018                                     3rd Qu.:   6929
## Max.   :2022                                     Max.   :31878711
```

Valores únicos de CATEGORIAS

```
cat("El número de valores únicos de PRODUCTOS es: ", length(unique(x = data$PRODUCTOS)), "\n")
```

```
## El número de valores únicos de PRODUCTOS es: 791
```

```
cat("El número de valores únicos de AÑO es: ", length(unique(x = data$AÑO)), "\n")
```

```
## El número de valores únicos de AÑO es: 23
```

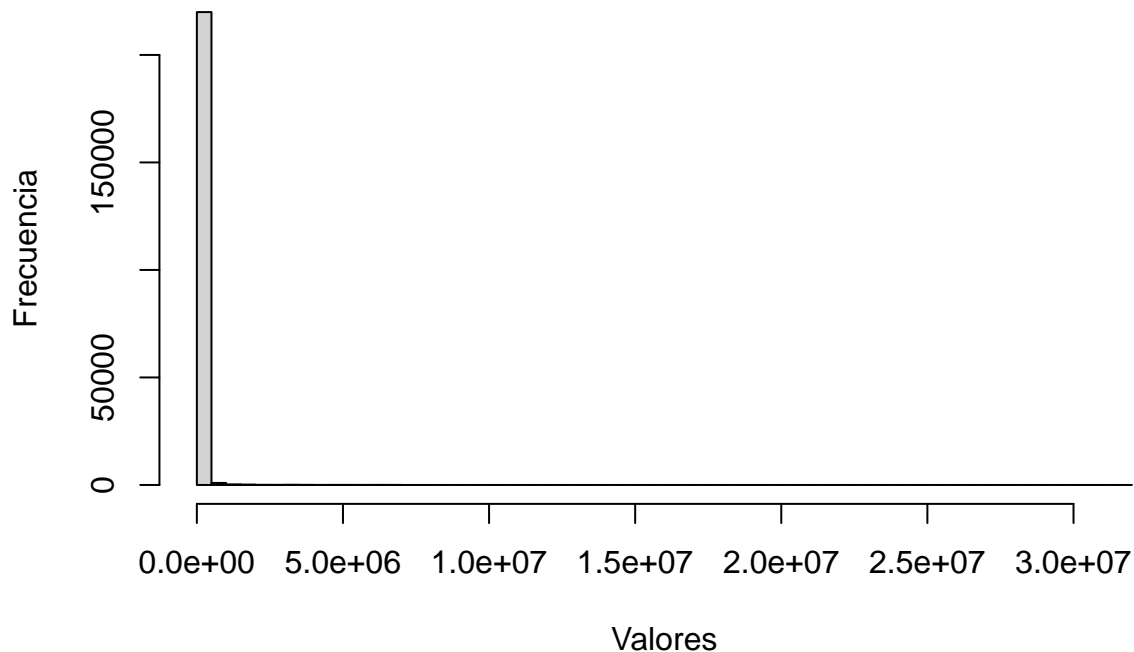
```
cat("El número de valores únicos de REGIONES es: ", length(unique(x = data$REGIONES)), "\n")
```

```
## El número de valores únicos de REGIONES es: 18
```

Acontinuación realizo un histograma de la variable VOLUMEN, para ver como se distribuyen sus valores

```
hist(  
  x = data$VOLUMEN,  
  main = "Histograma de VOLUMEN",  
  xlab = "Valores",  
  ylab = "Frecuencia",  
  breaks = 100,  
  axes = TRUE  
)
```

Histograma de VOLUMEN



Según los resultados obtenidos, las principales características del conjunto de datos son las siguientes:

- El conjunto de datos tiene 221,706 registros y 4 variables.
- Las variables son:
 - AÑO: contiene valores numéricos que representan años. El rango temporal abarca desde el año 2000 hasta el 2022, en total, 23 años diferentes.
 - PRODUCTOS: contiene valores de tipo cadena de texto, que representa los tipos de alimentos. Por lo que parece existen 791 tipos de alimentos únicos.
 - REGIONES: es otra variable, que también admite valores de tipo carácter, y que representa las diferentes regiones geográficas, esto es, las Comunidades Autónomas y España, en total 18 regiones.

- VOLUMEN: contiene valores numéricos y representa la cantidad de alimentos que se consumen por los hogares españoles. En este caso, nuestra variable respuesta es VOLUMEN ya que la pregunta a responder sería como varía la cantidad de alimento consumida por los hogares españoles por año, región y tipo de alimento.
- Los valores de la variable VOLUMEN presentan una distribución sesgada hacia la izquierda, esto quiere decir que la mayoría de sus valores se sitúan cerca del cero.

4. Ajuste de variables

Tras el análisis descriptivo del conjunto de datos, conviene ajustar algunas de las variables.

En este caso, la variable AÑO, que es numérica, no voy a modificarla, ya que esto me podría ayudar a realizar alguna operación matemática en el futuro.

Con respecto a las variables REGIONES y PRODUCTOS voy a factorizarla.

```
data$REGIONES <- as.factor(data$REGIONES)
data$PRODUCTOS <- as.factor(data$PRODUCTOS)
```

5. Detección y tratamiento de datos ausentes

Compruebo si existen valores ausentes en el conjunto de los datos.

```
any(is.na(data))
```

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

Según se puede apreciar, en el conjunto de los datos no existe valores ausentes.

6. Identificación de datos atípicos

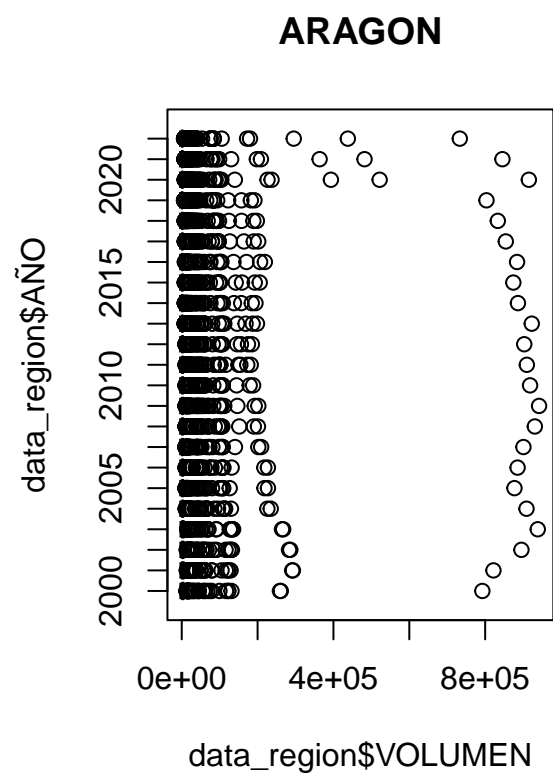
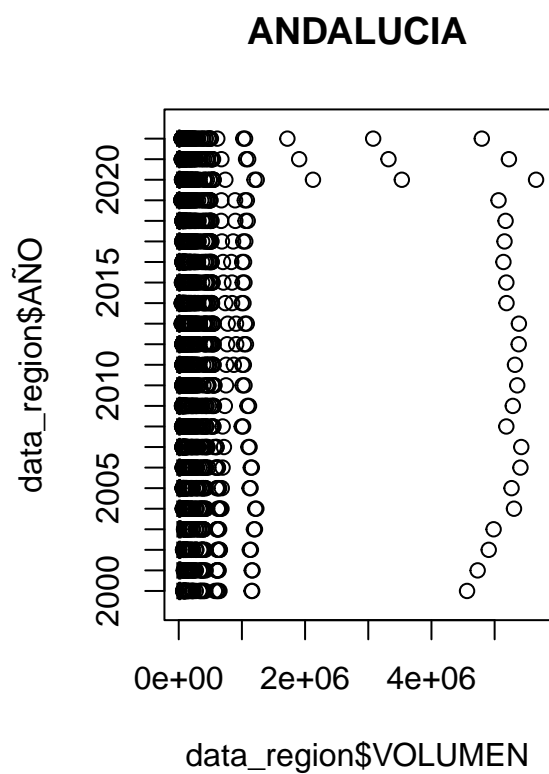
Mediante un gráfico boxplot, compruebo si existen valores atípicos en VOLUMEN por región y año.

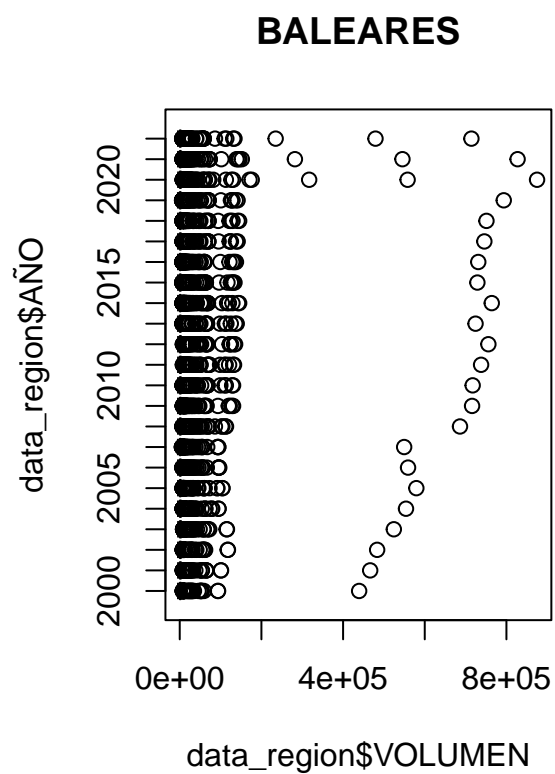
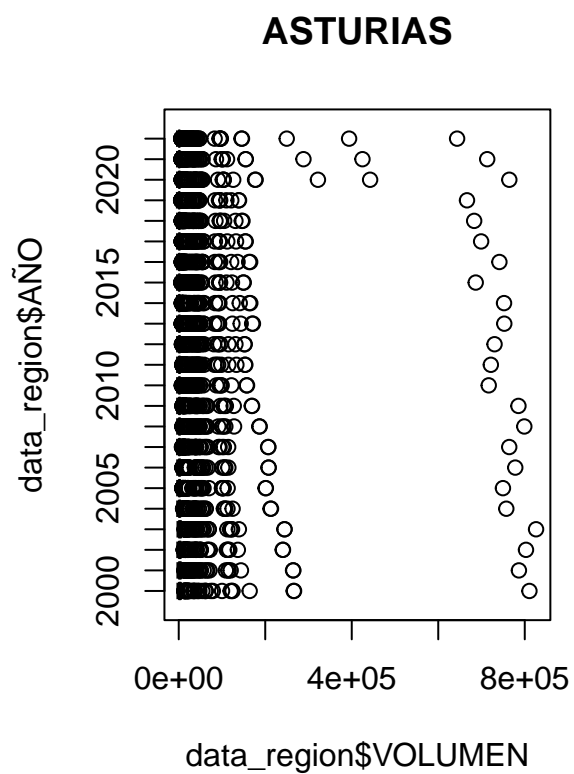
```
# Configurar la disposición de los gráficos (1 fila, 2 columnas)
par(mfrow = c(1, 2))

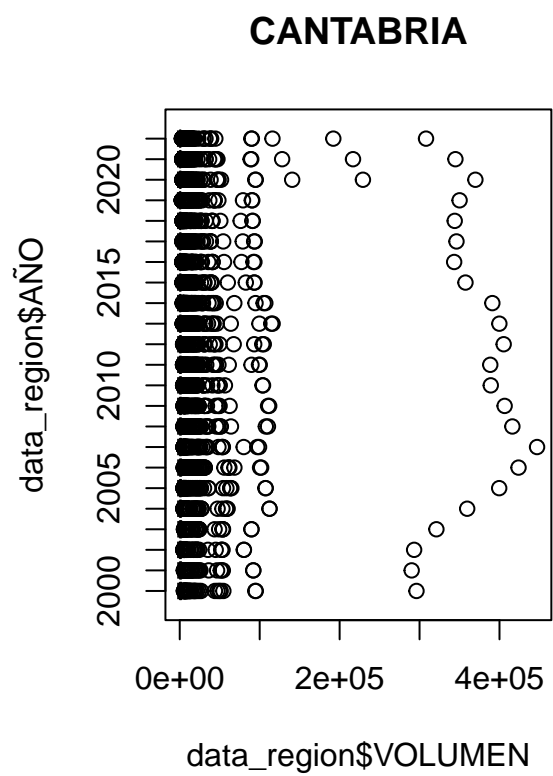
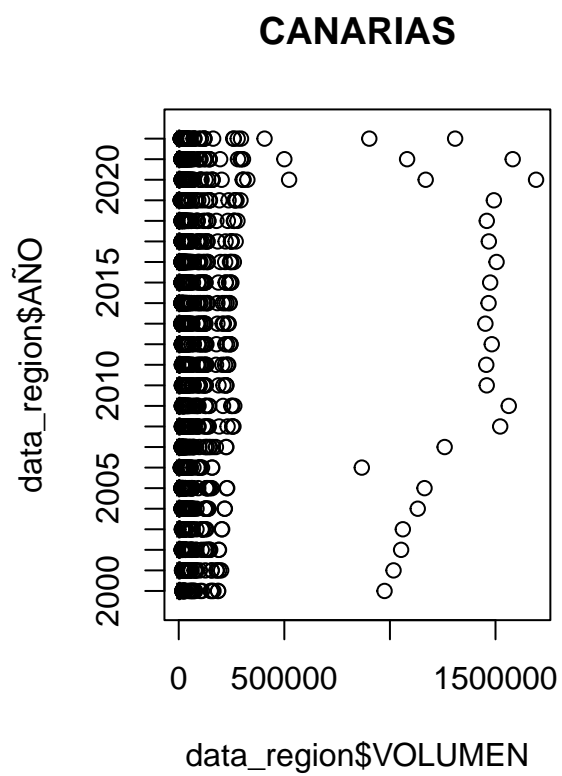
# Crear una lista de regiones únicas en tus datos
regiones_unicas <- unique(data$REGIONES)

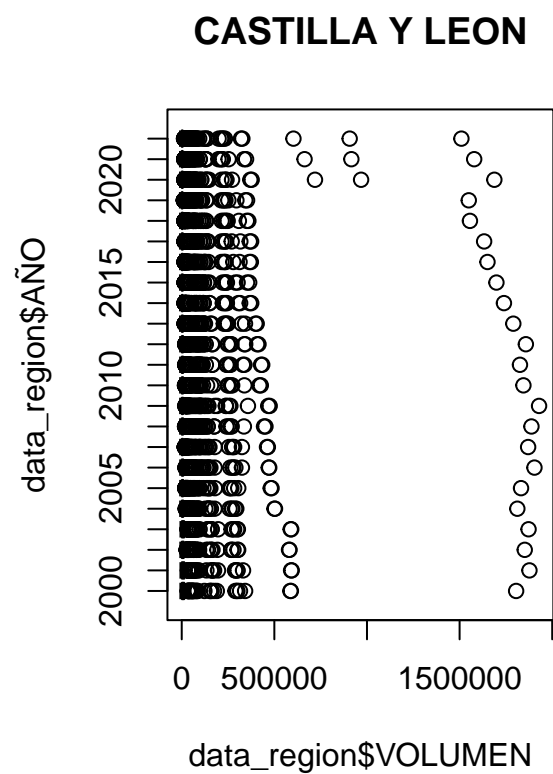
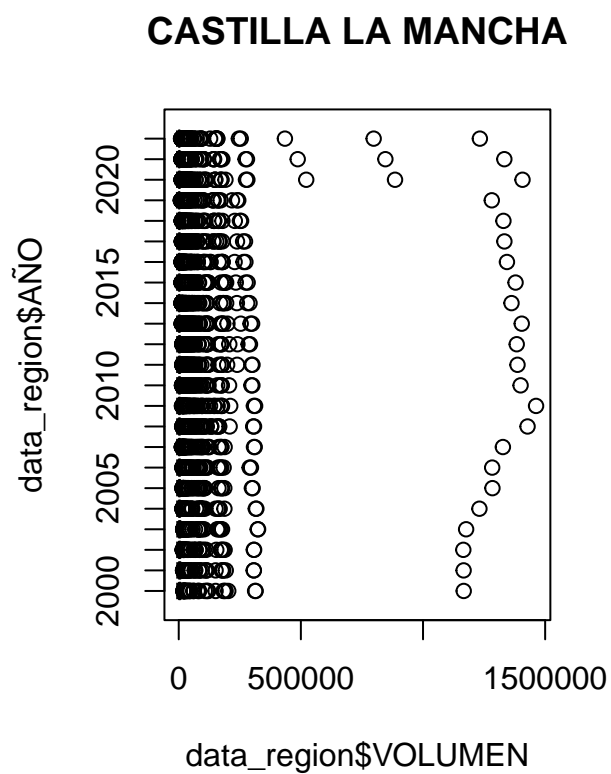
# Crear un bucle para generar boxplots por región y año
for (region in regiones_unicas) {
  # Filtrar los datos para la región actual
  data_region <- subset(data, REGIONES == region)

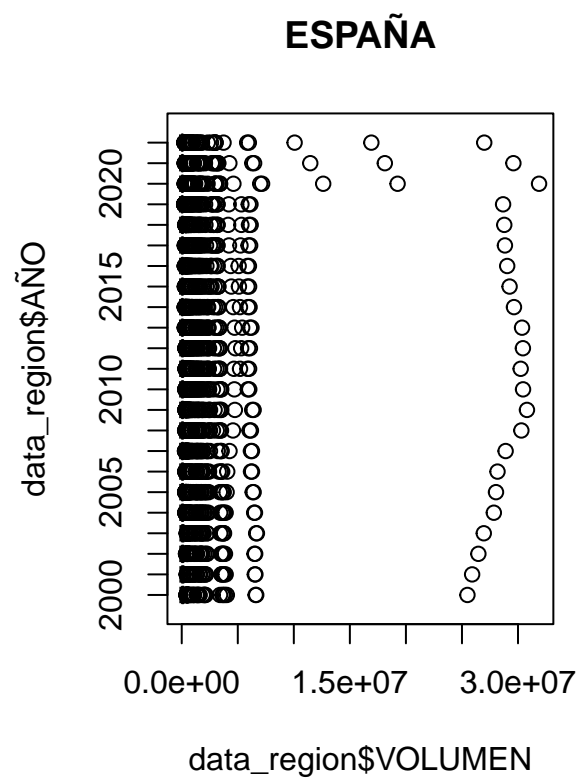
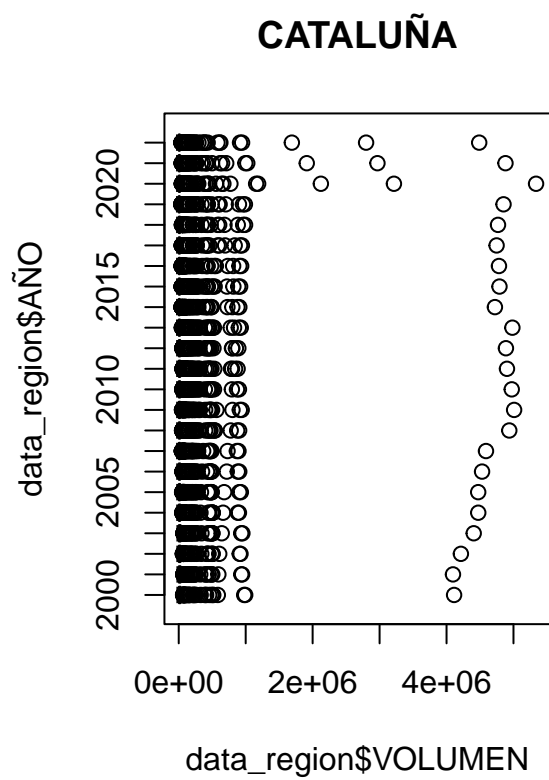
  # Crear el boxplot para la región y el año actual
  boxplot(
    data_region$VOLUMEN ~ data_region$AÑO,
    horizontal = TRUE,
    main = paste(region))
}
```

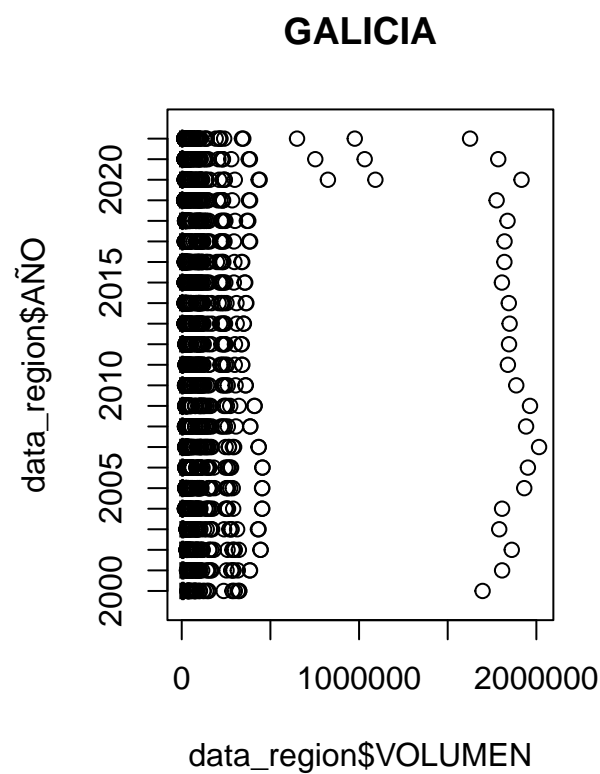
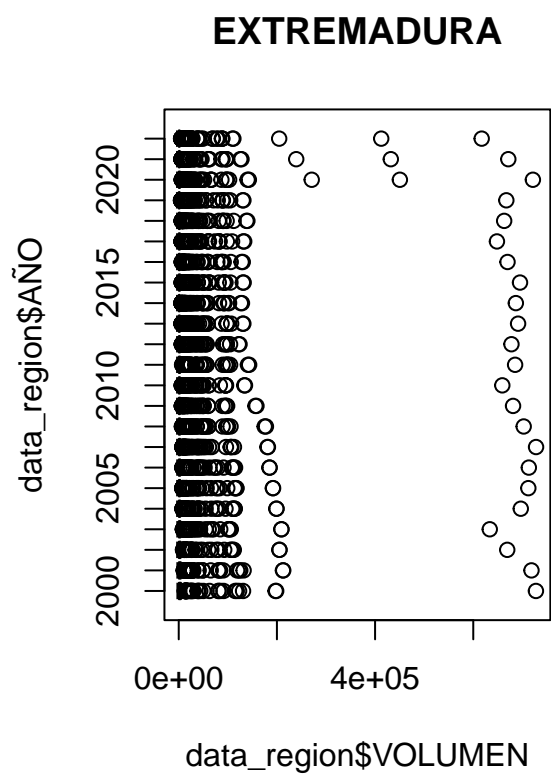


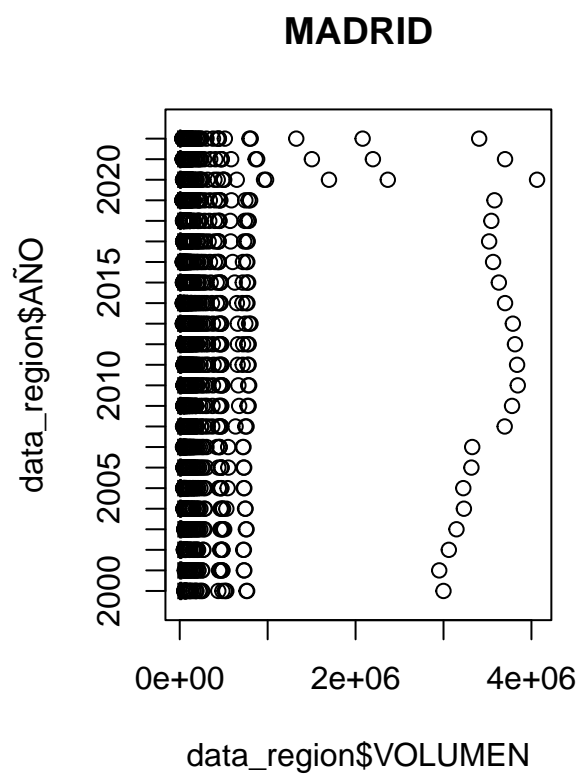
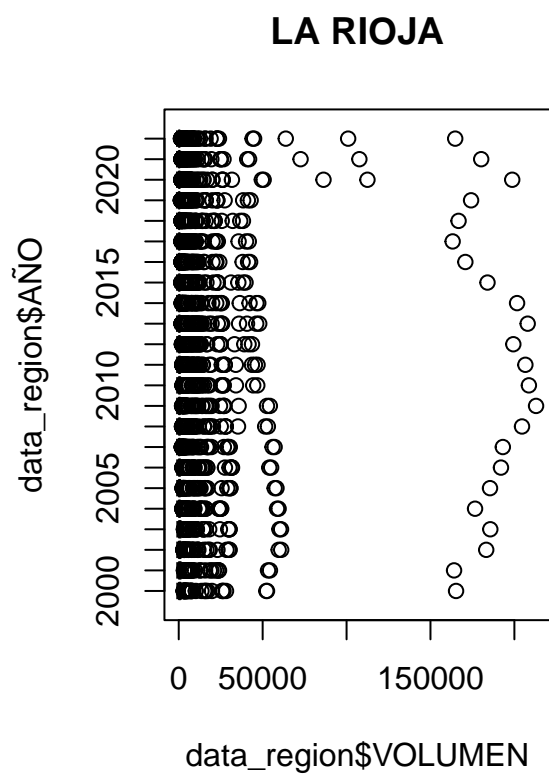


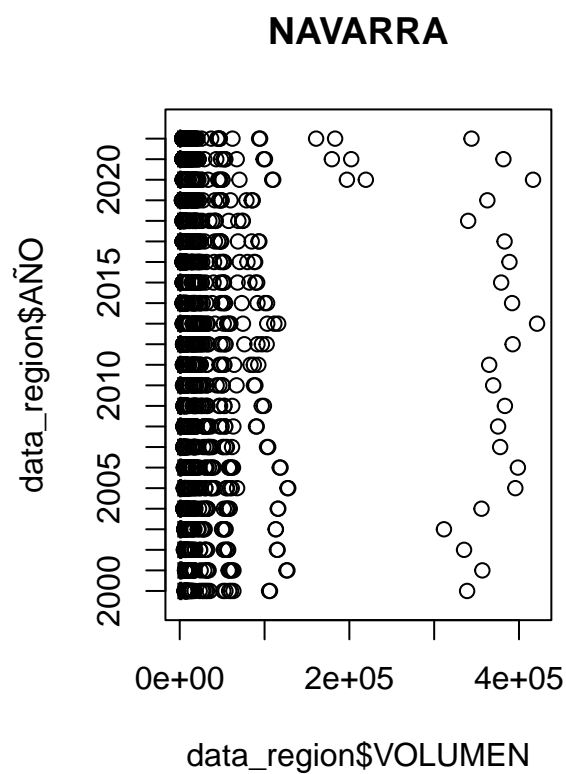
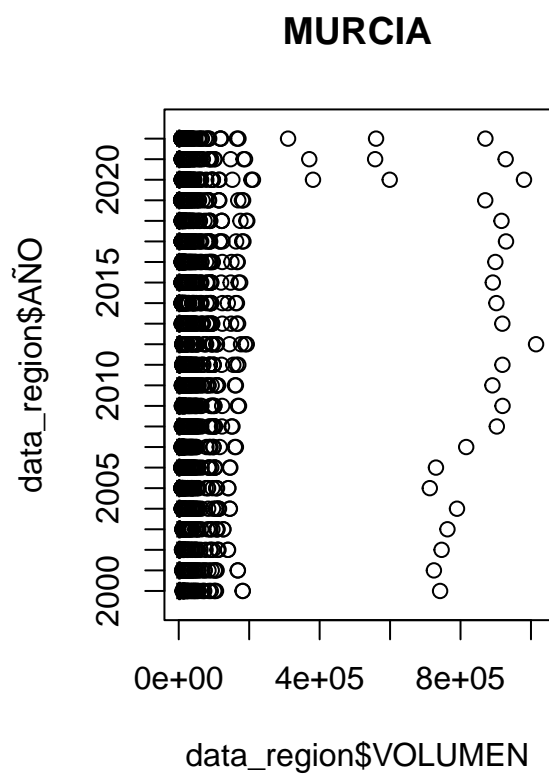


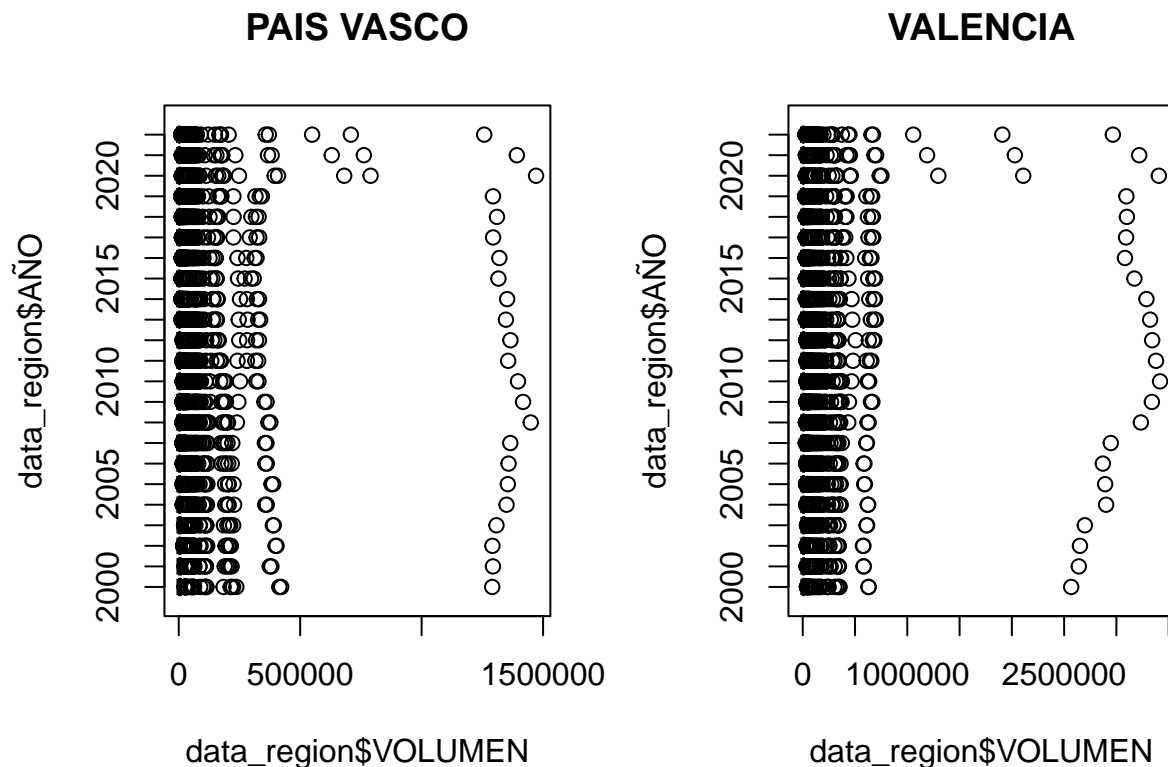










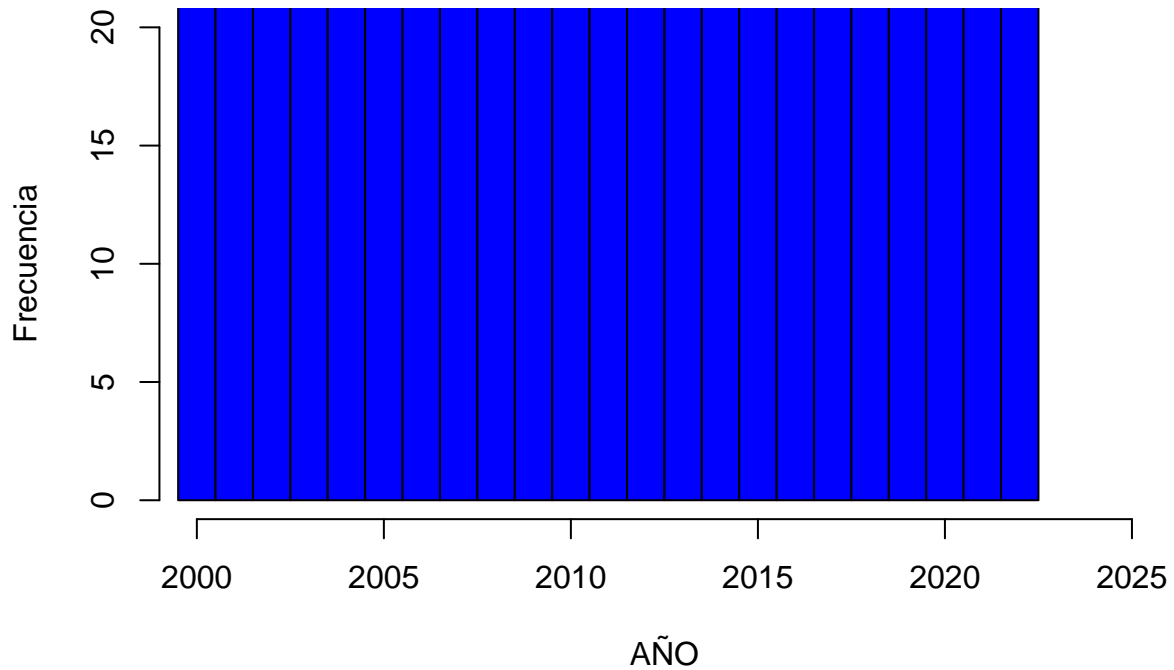


En conclusión, al examinar los valores atípicos en los gráficos de boxplot de “VOLUMEN” por región y año, hemos observado la presencia de variabilidad significativa en el consumo de alimentos a lo largo del tiempo y en diferentes regiones geográficas, lo cual era de esperar debido a las diferencias en las preferencias alimenticias y las influencias socioeconómicas. Si bien es cierto que hemos identificado valores que se desvían de la norma, es importante destacar que no todos los valores atípicos son necesariamente errores o anomalías; muchos de ellos reflejan patrones genuinos en los datos. La interpretación adecuada de estos valores atípicos requiere un conocimiento profundo del contexto y el dominio, y puede proporcionar información valiosa para entender las tendencias y las diferencias regionales en el consumo de alimentos a lo largo del tiempo. Por lo tanto, en lugar de descartar automáticamente los valores atípicos, debemos considerarlos como una parte esencial de la riqueza de nuestros datos y explorar sus implicaciones en el análisis.

A continuación, identifico los valores atípicos de las variables AÑO, REGIONES y CATEGORIAS.

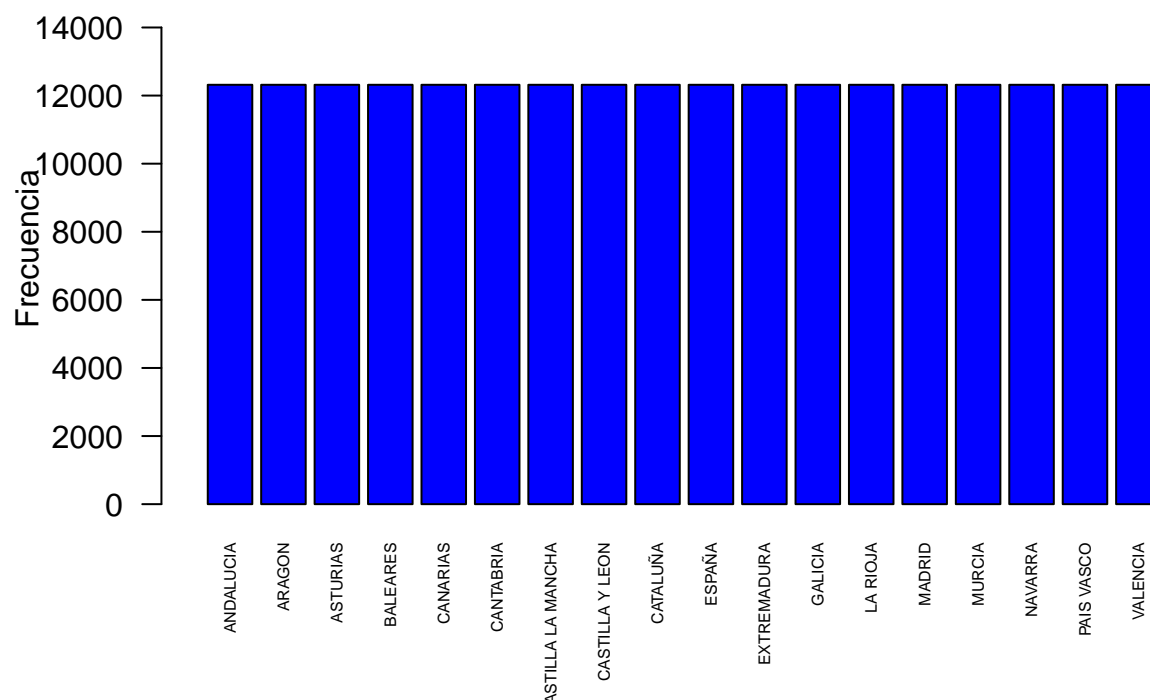
```
# Crear un histograma de la variable "AÑO"
hist(data$AÑO,
      main = "Histograma de AÑO",
      xlab = "AÑO",
      ylab = "Frecuencia",
      col = "blue",
      breaks = seq(min(data$AÑO) - 0.5, max(data$AÑO) + 0.5, by = 1),
      xlim = c(2000, 2025), # Ajustar límites del eje x
      ylim = c(0, 20)) # Ajustar límites del eje y
```

Histograma de AÑO



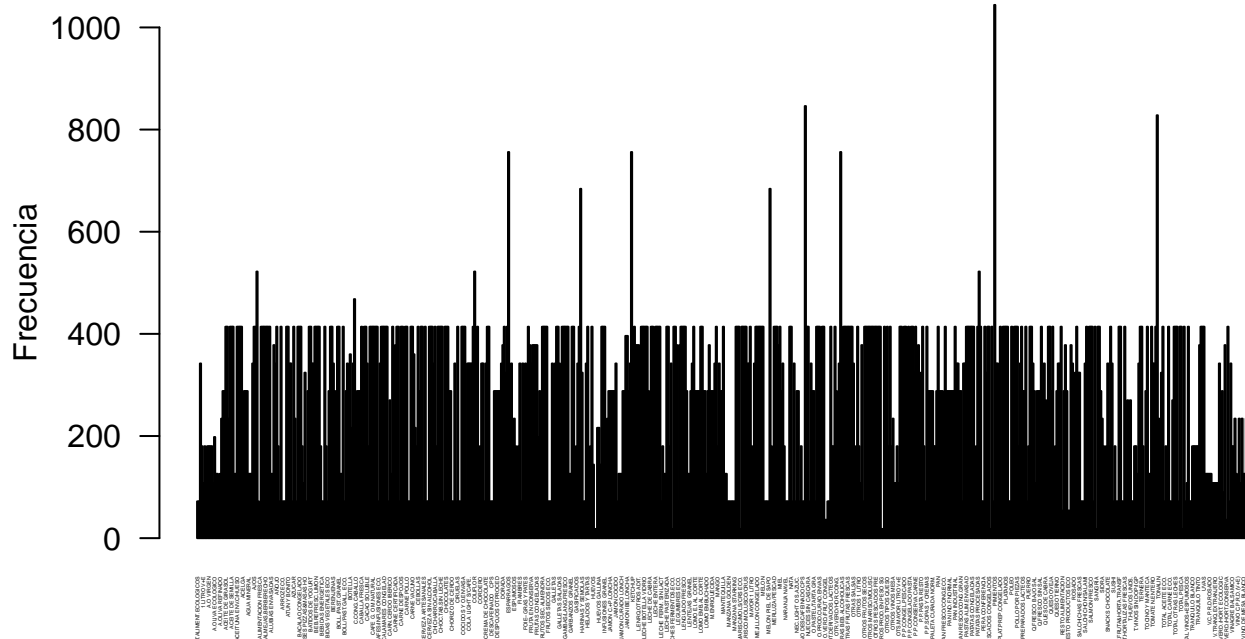
```
# Crear un gráfico de barras para la variable "REGIONES" con etiquetas personalizadas
barplot(table(data$REGIONES),
        main = "Histograma de REGIONES",
        ylab = "Frecuencia",
        col = "blue",
        border = "black",
        las = 2, # Rotar etiquetas en el eje x
        cex.names = 0.5, # Tamaño de las etiquetas de las regiones
        xlim = c(0, length(unique(data$REGIONES)) + 3), # Ajustar límites del eje x
        ylim = c(0, 15000)) # Ajustar límites del eje y
```

Histograma de REGIONES



```
# Crear un gráfico de barras para la variable "REGIONES" con etiquetas personalizadas
barplot(table(data$PRODUCTOS),
  main = "Histograma de PRODUCTOS",
  ylab = "Frecuencia",
  col = "blue",
  border = "black",
  las = 2, # Rotar etiquetas en el eje x
  cex.names = 0.2, # Tamaño de las etiquetas de las regiones
  xlim = c(0, length(unique(data$PRODUCTOS)) + 3), # Ajustar límites del eje x
  ylim = c(0, 1000)) # Ajustar límites del eje y
```

Histograma de PRODUCTOS



Trasa realizar el análisis de los gráficos boxplot e histogramas, he llegado a las siguientes conclusiones:

- Los valores únicos de las variables REGIONES y AÑOS se repiten el mismo número de veces cada uno de ellos.
- Con respecto a los valores de la variable CATEGORIAS, por regla general se repiten el mismo número de veces, a excepción de unos cuantos, pero esto se debe a modificaciones en la metodología en la toma de los datos, de tal forma, que algunas categorías de alimentos se empezaron a contabilizar en unos años y otros en otros.
- Con respecto a los valores atípicos en los gráficos de boxplot de “VOLUMEN” por región y año, he observado la presencia de variabilidad significativa en el consumo de alimentos a lo largo del tiempo y en diferentes regiones geográficas, lo cual era de esperar debido a las diferencias en las preferencias alimenticias. Si bien es cierto que hemos identificado valores que se desvían de la norma, es importante destacar que no todos los valores atípicos son necesariamente errores o anomalías; muchos de ellos reflejan patrones genuinos en los datos. La interpretación adecuada de estos valores atípicos requiere un conocimiento profundo del contexto y el dominio, y puede proporcionar información valiosa para entender las tendencias y las diferencias regionales en el consumo de alimentos a lo largo del tiempo. Por lo tanto, en lugar de descartar automáticamente los valores atípicos, debemos considerarlos como una parte esencial de la riqueza de nuestros datos y explorar sus implicaciones en el análisis.

8. Conclusiones

El análisis exploratorio de datos se llevó a cabo en un conjunto de datos que abarca el volumen de alimentos consumidos por hogares españoles durante el período de 2000 a 2022. Estos datos se extrajeron de 22 archivos Excel, uno para cada año, disponibles en un repositorio de GitHub. Durante el EDA, se obtuvieron las siguientes conclusiones:

En una exploración inicial de los diferentes dataframe, e observó una variabilidad significativa en la cantidad de variables entre los distintos dataframes. Algunos contenían 19 variables, mientras que otros presentaban 27 debido a cambios en la metodología a partir de 2004 y en el 2019. Se redujo el número de variables, conservando solo las que corresponden a España y las Comunidades Autónomas.

De la misma forma, dependiendo de la metodología adoptada, el número de registros, e incluso su denominación, varían según los años. La selección de estos registros dependerán en medida de los objetivos del análisis. Así, si se pretende un análisis general, se podrá seleccionar aquellas categorías, como TOTAL CARNE, TOTAL PESCA, etc, que faciliten ese fin. Si se quiere un análisis más detallado, se podrá seleccionar productos más concreto, como carne de vaca congelada, etc.

Una vez transformado los datos, se unieron todos los dataframe en uno sólo, dando lugar a un nuevo conjunto dedatos cuyas principales características son la siguientes:

- El conjunto de datos tiene 221,706 registros y 4 variables.
- Las variables son:
 - **AÑO**: contine valores numéricos que representan años. El rango temporal abarca desde el año 2000 hasta el 2022, en total, 23 años diferentes.
 - **PRODUCTOS**: contiene valores de tipo cadena de texto, que representa los tipos de alimentos. Existen 791 tipos de alimentos únicos.
 - **REGIONES**: es otra variable, que también admite valores de tipo caracter, y que representa las diferentes regiones geográficas, esto es, las Comunidades Autónomas y España, en total 18 regiones.
 - **VOLUMEN**: contiene valores numéricos y representa la cantidad de alimentos que se consumen por los hogares españoles. En este caso, nuestra variable respuesta es VOLUMEN ya que la pregunta a responder sería como varía la cantidad de alimento consumida por los hogares españoles por año, región y tipo de alimento.
- Los valores de la variable VOLUMEN presentan una distribución sesgada hacia la izquierda, esto quiere decir que la mayoría de sus valores se sitúan cerca del cero.
- **Datos Completos**: No se encontraron datos faltantes en el conjunto de datos resultante.
- **Datos Atípicos**: Los gráficos de boxplot de “VOLUMEN” por región y año revelaron una variabilidad significativa en el consumo de alimentos a lo largo del tiempo y en diferentes regiones geográficas. Si bien se identificaron valores atípicos, estos reflejan patrones genuinos en los datos y deben considerarse como parte integral de la riqueza de la información.

En resumen, el análisis exploratorio de datos proporciona una sólida base para futuras investigaciones relacionadas con el consumo de alimentos en hogares españoles. Se enfatiza la importancia de considerar la variabilidad y los cambios en los datos a lo largo del tiempo y se destaca la necesidad de mantener registros completos.