

Proyecto Módulo 2

Ignacio Mondragon Moreno

2022-06-09

% !TEX encoding = UTF-8 Unicode

carga de fichero local

```
product = read.csv("../DS/productos-verano-calificación-rendimiento_2020-08.csv")
cat = read.csv("../DS/categorias-unicas.csv")
catOrd = read.csv("../DS/categorias-unicas-ordenadas-recuento.csv")
```

Filas y columnas DF

```
# Primeras 20 filas
head(product,2)
```

```
##
## 1 2020 Summer Vintage Flamingo Print Pajamas Set Casual Loose T Shirt Top And Elastic Shorts Women S
## 2                                     Women's Casual Summer Sleeveless Sexy Mi
##                               Titulo_corto precio precio_al_por_menor
## 1 2020 Summer Vintage Flamingo Print Paja      16              14
## 2 Women's Casual Summer Sleeveless Sexy Mi      8              22
## comprador_de_divisas unidades_vendidas utiliza_refuerzos_publicitarios
## 1 EUR              100              0
## 2 EUR              20000             1
## clasificacion conteo_de_calificacion clasificacion_cinco_cuenta
## 1 3.76              54              26
## 2 3.45              6135             2269
## clasificacion_cuatro_cuenta clasificacion_de_tres_cuentas
## 1 8              10
## 2 1027             1118
## clasificacion_de_dos_cuentas calificar_un_conteo insignias_cuentan
## 1 1              9              0
## 2 644             1077             0
## insignia_producto_local insignia_de_calidad_del_producto
## 1 0              0
## 2 0              0
## placa_de_envio_rápido
## 1 0
```

```

## 2          0
##
## 1 Summer,Fashion,womenunderwearsuit,printedpajamasset,womencasualshort,Women's Fashion,flamingo,loung
## 2
##   color_del_producto ID_de_tamano_de_variacion_de_producto
## 1           white                      M
## 2           green                      XS
##   inventario_de_variaciones_de_productos nombre_de_la_opcion_de_envio
## 1                      50          Livraison standard
## 2                      50          Livraison standard
##   precio_de_opcion_de_envio el_envio_es_express paises_enviados_a
## 1                      4                      0          34
## 2                      2                      0          41
##   inventario_total tiene_banner_de_urgencia texto_de_urgencia pais_de_origen
## 1                      50          1 Quantité limitée !          CN
## 2                      50          1 Quantité limitée !          CN
##   titulo_de_comerciante nombre_del_comerciante
## 1           zgrdejia          zgrdejia
## 2           SaraHouse          sarahouse
##   subtitulo_de_informacion_comercial conteo_de_calificaciones_del_comerciante
## 1                      (568 notes)          568
## 2 83 % avis positifs (17,752 notes)          17752
##   calificacion_del_comerciante comerciante_id
## 1          4.128521 595097d6a26f6e070cb878d1
## 2          3.899673 56458aa03a698c35c9050988
##   el_comerciante_tiene_una_foto_de_perfil foto_de_perfil_del_comerciante
## 1                      0
## 2                      0
##   URL_del_producto
## 1 https://www.wish.com/c/5e9ae51d43d6a96e303acdb0
## 2 https://www.wish.com/c/58940d436a0d3d5da4e95a38
##   imagen_del_producto
## 1 https://contesting.wish.com/api/webimage/5e9ae51d43d6a96e303acdb0-medium.jpg
## 2 https://contesting.wish.com/api/webimage/58940d436a0d3d5da4e95a38-medium.jpg
##   ID_del_producto tema mes_de_rastreo
## 1 5e9ae51d43d6a96e303acdb0 summer          2020-08
## 2 58940d436a0d3d5da4e95a38 summer          2020-08

```

```

# Ultimas 2 filas
tail(product,2)

```

```

##                                     titulo_origen
## 1572 Fashion Women Summer Sleeveless Dresses V Neck Backless Lace Dress Ladies Beach Dresses White D
## 1573                                     Fashion Women Yoga Pants Slim Fit Fitness Running Legg
##   Titulo_corto precio precio_al_por_menor
## 1572 Fashion Women Summer Sleeveless Dresses          13          11
## 1573 Fashion Women Yoga Pants Slim Fit Fitnes          7          6
##   comprador_de_divisas unidades_vendidas utiliza_refuerzos_publicitarios
## 1572          EUR          100          0
## 1573          EUR          100          1
##   clasificacion conteo_de_calificacion calificacion_cinco_cuenta
## 1572          2.50          2          0
## 1573          4.07          14          8
##   calificacion_cuatro_cuenta calificacion_de_tres_cuentas

```

```

## 1572          1          0
## 1573          3          1
##      calificacion_de_dos_cuentas calificar_un_conteo insignias_cuentan
## 1572          0          1          0
## 1573          0          2          0
##      insignia_producto_local insignia_de_calidad_del_producto
## 1572          0          0
## 1573          0          0
##      placa_de_envio_rápido
## 1572          0
## 1573          0
##
## 1572 Summer,fashion women,Fashion,Lace,Dresses,Dress,Lace Dress,Women's Fashion,ladies dress,beach d
## 1573
##      color_del_producto ID_de_tamano_de_variacion_de_producto
## 1572          white          Size S.
## 1573          red          S
##      inventario_de_variaciones_de_productos nombre_de_la_opcion_de_envio
## 1572          36          Livraison standard
## 1573          50          Livraison standard
##      precio_de_opcion_de_envio el_envio_es_express paises_enviados_a
## 1572          3          0          29
## 1573          2          0          41
##      inventario_total tiene_banner_de_urgencia texto_de_urgencia pais_de_origen
## 1572          50          NA          CN
## 1573          50          NA          CN
##      titulo_de_comerciante nombre_del_comerciante
## 1572          modai          modai
## 1573          AISHOPPINGMALL          aishoppingmall
##      subtitulo_de_informacion_comercial
## 1572          77 % avis positifs (31 notes)
## 1573          90 % avis positifs (7,023 notes)
##      conteo_de_calificaciones_del_comerciante calificacion_del_comerciante
## 1572          31          3.774194
## 1573          7023          4.235939
##      comerciante_id el_comerciante_tiene_una_foto_de_perfil
## 1572 5d56b32c40defd78043d5af9          0
## 1573 5a409cf87b584e7951b2e25f          0
##      foto_de_perfil_del_comerciante
## 1572
## 1573
##      URL_del_producto
## 1572 https://www.wish.com/c/5eda07ab0e295c2097c36590
## 1573 https://www.wish.com/c/5e857321f53c3d2d8f25e7ed
##      imagen_del_producto
## 1572 https://contestimg.wish.com/api/webimage/5eda07ab0e295c2097c36590-medium.jpg
## 1573 https://contestimg.wish.com/api/webimage/5e857321f53c3d2d8f25e7ed-medium.jpg
##      ID_del_producto tema mes_de_rastreo
## 1572 5eda07ab0e295c2097c36590 summer          2020-08
## 1573 5e857321f53c3d2d8f25e7ed summer          2020-08

```

```

#Estructura del Data Frame
str(product)

```

```
## 'data.frame': 1573 obs. of 43 variables:
## $ titulo_original : chr "2020 Summer Vintage Flamingo Print Pajamas Set C
## $ Titulo_corto : chr "2020 Summer Vintage Flamingo Print Paja" "Women's
## $ precio : num 16 8 8 8 2.72 3.92 7 12 11 5.78 ...
## $ precio_al_por_menor : int 14 22 43 8 3 9 6 11 84 22 ...
## $ comprador_de_divisas : chr "EUR" "EUR" "EUR" "EUR" ...
## $ unidades_vendidas : int 100 20000 100 5000 100 10 50000 1000 100 5000 ...
## $ utiliza_refuerzos_publicitarios : int 0 1 0 1 1 0 0 0 1 0 ...
## $ clasificacion : num 3.76 3.45 3.57 4.03 3.1 5 3.84 3.76 3.47 3.6 ...
## $ conteo_de_calificacion : int 54 6135 14 579 20 1 6742 286 15 687 ...
## $ calificacion_cinco_cuenta : int 26 2269 5 295 6 1 3172 120 6 287 ...
## $ calificacion_cuatro_cuenta : int 8 1027 4 119 4 0 1352 56 2 128 ...
## $ calificacion_de_tres_cuentas : int 10 1118 2 87 2 0 971 61 3 92 ...
## $ calificacion_de_dos_cuentas : int 1 644 0 42 2 0 490 18 1 68 ...
## $ calificar_un_conteo : int 9 1077 3 36 6 0 757 31 3 112 ...
## $ insignias_cuentan : int 0 0 0 0 0 0 0 0 0 0 ...
## $ insignia_producto_local : int 0 0 0 0 0 0 0 0 0 0 ...
## $ insignia_de_calidad_del_producto : int 0 0 0 0 0 0 0 0 0 0 ...
## $ placa_de_envio_rápido : int 0 0 0 0 0 0 0 0 0 0 ...
## $ etiquetas : chr "Summer,Fashion,womenunderwearsuit,printedpajamass
## $ color_del_producto : chr "white" "green" "leopardprint" "black" ...
## $ ID_de_tamano_de_variacion_de_producto : chr "M" "XS" "XS" "M" ...
## $ inventario_de_variaciones_de_productos : int 50 50 1 50 1 1 50 50 50 50 ...
## $ nombre_de_la_opcion_de_envio : chr "Livraison standard" "Livraison standard" "Livrais
## $ precio_de_opcion_de_envio : int 4 2 3 2 1 1 2 3 2 2 ...
## $ el_envio_es_express : int 0 0 0 0 0 0 0 0 0 0 ...
## $ paises_enviados_a : int 34 41 36 41 35 40 31 139 36 33 ...
## $ inventario_total : int 50 50 50 50 50 50 50 50 50 50 ...
## $ tiene_banner_de_urgencia : int 1 1 1 NA 1 NA NA NA 1 NA ...
## $ texto_de_urgencia : chr "Quantité limitée !" "Quantité limitée !" "Quantit
## $ pais_de_origen : chr "CN" "CN" "CN" "CN" ...
## $ titulo_de_comerciante : chr "zgrdejia" "SaraHouse" "hxt520" "allenfan" ...
## $ nombre_del_comerciante : chr "zgrdejia" "sarahouse" "hxt520" "allenfan" ...
## $ subtítulo_de_informacion_comercial : chr "(568 notes)" "83 % avis positifs (17,752 notes)"
## $ conteo_de_calificaciones_del_comerciante : int 568 17752 295 23832 14482 65 10194 342 330 5534 ..
## $ calificacion_del_comerciante : num 4.13 3.9 3.99 4.02 4 ...
## $ comerciante_id : chr "595097d6a26f6e070cb878d1" "56458aa03a698c35c90509
## $ el_comerciante_tiene_una_foto_de_perfil : int 0 0 0 0 0 0 1 0 0 0 ...
## $ foto_de_perfil_del_comerciante : chr "" "" "" "" ...
## $ URL_del_producto : chr "https://www.wish.com/c/5e9ae51d43d6a96e303acdb0"
## $ imagen_del_producto : chr "https://contestimg.wish.com/api/webimage/5e9ae51d
## $ ID_del_producto : chr "5e9ae51d43d6a96e303acdb0" "58940d436a0d3d5da4e95a
## $ tema : chr "summer" "summer" "summer" "summer" ...
## $ mes_de_rastreo : chr "2020-08" "2020-08" "2020-08" "2020-08" ...
```

```
# Nombres de columnas -> colnames -> rownames ->dimnames"trae filas y columnas"
names(product)
```

```
## [1] "titulo_original"
## [2] "Titulo_corto"
## [3] "precio"
## [4] "precio_al_por_menor"
## [5] "comprador_de_divisas"
## [6] "unidades_vendidas"
```

```
## [7] "utiliza_refuerzos_publicitarios"
## [8] "clasificacion"
## [9] "conteo_de_calificacion"
## [10] "calificacion_cinco_cuenta"
## [11] "calificacion_cuatro_cuenta"
## [12] "calificacion_de_tres_cuentas"
## [13] "calificacion_de_dos_cuentas"
## [14] "calificar_un_conteo"
## [15] "insignias_cuentan"
## [16] "insignia_producto_local"
## [17] "insignia_de_calidad_del_producto"
## [18] "placa_de_envio_rápido"
## [19] "etiquetas"
## [20] "color_del_producto"
## [21] "ID_de_tamano_de_variacion_de_producto"
## [22] "inventario_de_variaciones_de_productos"
## [23] "nombre_de_la_opcion_de_envio"
## [24] "precio_de_opcion_de_envio"
## [25] "el_envio_es_express"
## [26] "países_enviados_a"
## [27] "inventario_total"
## [28] "tiene_banner_de_urgencia"
## [29] "texto_de_urgencia"
## [30] "país_de_origen"
## [31] "titulo_de_comerciante"
## [32] "nombre_del_comerciante"
## [33] "subtitulo_de_informacion_comercial"
## [34] "conteo_de_calificaciones_del_comerciante"
## [35] "calificacion_del_comerciante"
## [36] "comerciante_id"
## [37] "el_comerciante_tiene_una_foto_de_perfil"
## [38] "foto_de_perfil_del_comerciante"
## [39] "URL_del_producto"
## [40] "imagen_del_producto"
## [41] "ID_del_producto"
## [42] "tema"
## [43] "mes_de_rastreo"
```

```
#Numero de filas -> ncol -> dim -> numero de filas y columnas
nrow(product)
```

```
## [1] 1573
```

```
#Extraer valores de una columna especifica, en un vector se puede especificar el inicio y termino por ejemplo:
```

Acceso al DF - Sub-DF

```
#Trae de la fila 1 a la 2 y todas las columnas -> Trae de la fila 1 a la 2 y columnas de la 2 a la 4 por ejemplo:
product[1:2,]
```

```
##
```

```

## 1 2020 Summer Vintage Flamingo Print Pajamas Set Casual Loose T Shirt Top And Elastic Shorts Women's
## 2 Women's Casual Summer Sleeveless Sexy Mi
## Titulo_corto precio precio_al_por_menor
## 1 2020 Summer Vintage Flamingo Print Paja 16 14
## 2 Women's Casual Summer Sleeveless Sexy Mi 8 22
## comprador_de_divisas unidades_vendidas utiliza_refuerzos_publicitarios
## 1 EUR 100 0
## 2 EUR 20000 1
## clasificacion conteo_de_calificacion calificacion_cinco_cuenta
## 1 3.76 54 26
## 2 3.45 6135 2269
## calificacion_cuatro_cuenta calificacion_de_tres_cuentas
## 1 8 10
## 2 1027 1118
## calificacion_de_dos_cuentas calificar_un_conteo insignias_cuentan
## 1 1 9 0
## 2 644 1077 0
## insignia_producto_local insignia_de_calidad_del_producto
## 1 0 0
## 2 0 0
## placa_de_envio_rápido
## 1 0
## 2 0
##
## 1 Summer,Fashion,womenunderwearsuit,printedpajamasset,womencasualshort,Women's Fashion,flamingo,loung
## 2
## color_del_producto ID_de_tamano_de_variacion_de_producto
## 1 white M
## 2 green XS
## inventario_de_variaciones_de_productos nombre_de_la_opcion_de_envio
## 1 50 Livraison standard
## 2 50 Livraison standard
## precio_de_opcion_de_envio el_envio_es_express paises_enviados_a
## 1 4 0 34
## 2 2 0 41
## inventario_total tiene_banner_de_urgencia texto_de_urgencia pais_de_origen
## 1 50 1 Quantité limitée ! CN
## 2 50 1 Quantité limitée ! CN
## titulo_de_comerciante nombre_del_comerciante
## 1 zgrdejia zgrdejia
## 2 SaraHouse sarahouse
## subtitulo_de_informacion_comercial conteo_de_calificaciones_del_comerciante
## 1 (568 notes) 568
## 2 83 % avis positifs (17,752 notes) 17752
## calificacion_del_comerciante comerciante_id
## 1 4.128521 595097d6a26f6e070cb878d1
## 2 3.899673 56458aa03a698c35c9050988
## el_comerciante_tiene_una_foto_de_perfil foto_de_perfil_del_comerciante
## 1 0
## 2 0
## URL_del_producto
## 1 https://www.wish.com/c/5e9ae51d43d6a96e303acdb0
## 2 https://www.wish.com/c/58940d436a0d3d5da4e95a38
## imagen_del_producto

```

```
## 1 https://contesting.wish.com/api/webimage/5e9ae51d43d6a96e303acdb0-medium.jpg
## 2 https://contesting.wish.com/api/webimage/58940d436a0d3d5da4e95a38-medium.jpg
##           ID_del_producto   tema mes_de_rastreo
## 1 5e9ae51d43d6a96e303acdb0 summer      2020-08
## 2 58940d436a0d3d5da4e95a38 summer      2020-08
```

```
#Filtrado df[df$Columna == "valor" & df$ColumnaInt>4,]
nrow(product[product$ID_de_tamano_de_variacion_de_producto == "M" & product$precio>4,])
```

```
## [1] 197
```

```
#Filtrado genera un sub DF-> Sobre ese solo traer la filas de 1 y 3 "c(1,3)" y columna c(2,5)
nrow(product[product$ID_de_tamano_de_variacion_de_producto == "M" & product$precio>4,][c(1,3),c(2,5)])
```

```
## [1] 2
```

Graficos basicas plot

```
conteoCalificaciones = product[['conteo_de_calificacion']]
paísesEnvio = product[['países_enviados_a']]
plot(paísesEnvio, conteoCalificaciones,
     xlab='País',
     ylab='Calificación 5',
     main='Escala lineal')
```

```
plot(paísesEnvio, conteoCalificaciones,
     log = 'x',
     xlab='País',
     ylab='Calificación 5',
     main='Escala logaritmica',
     pch=20,
     col = 'cyan3')
```

Histograma basico hist

```
hist(product$calificacion_del_comerciante,
     main = "Histograma de calificaciones",
     col = "cyan3",
     xlab = "Calificación",
     ylab = "Frecuencia")
```

Diagrama de frecuencias unidimensionales barplot

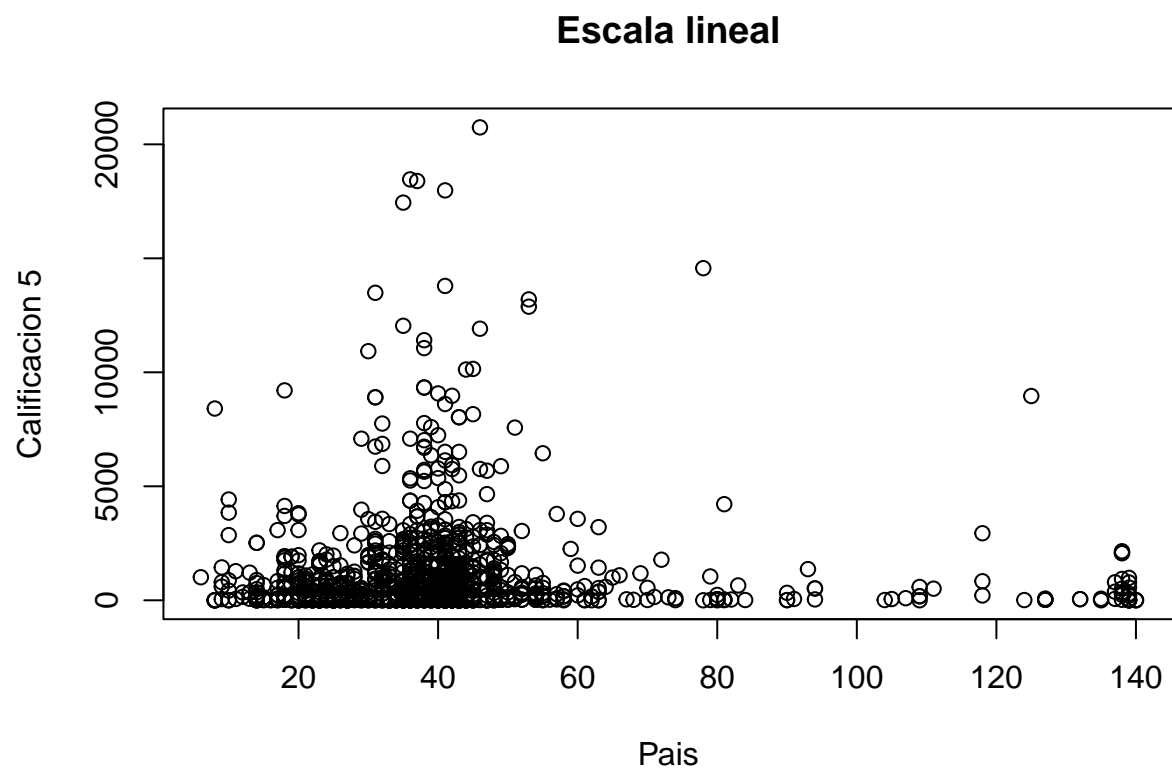


Figure 1: Grafico basico usando plot - calificaciones vs paises

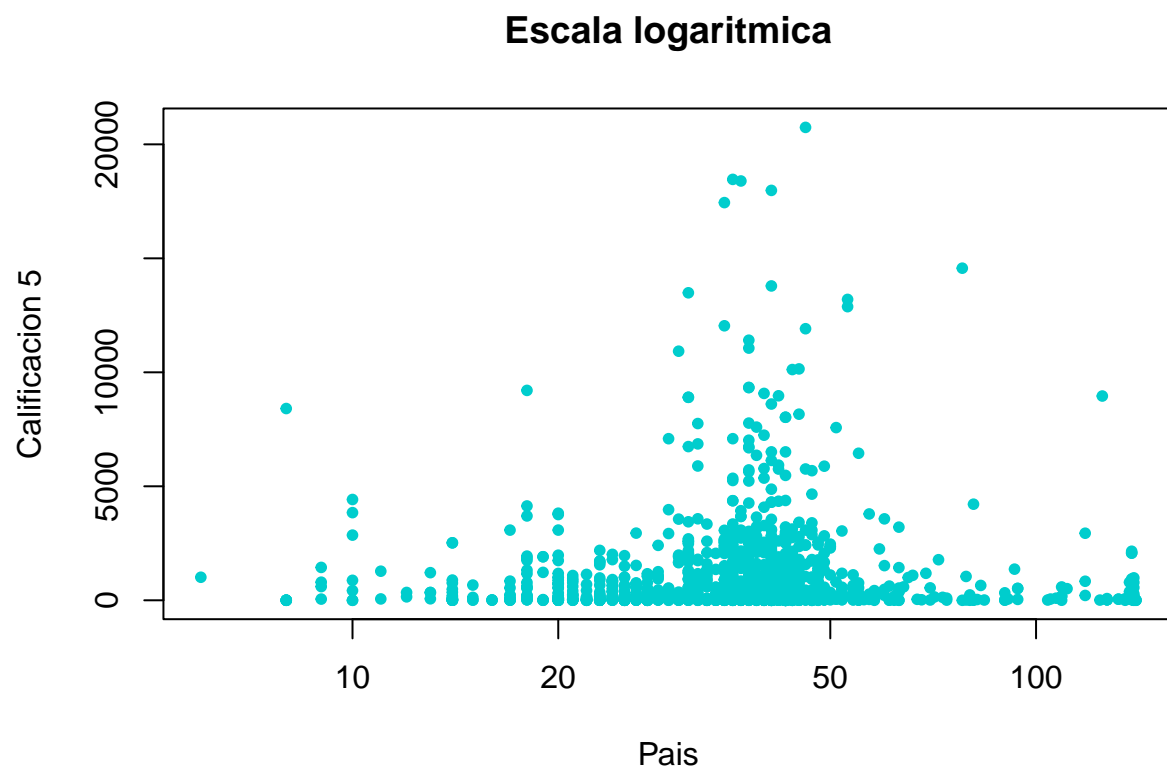


Figure 2: Grafico basico usando plot - calificaciones vs paises

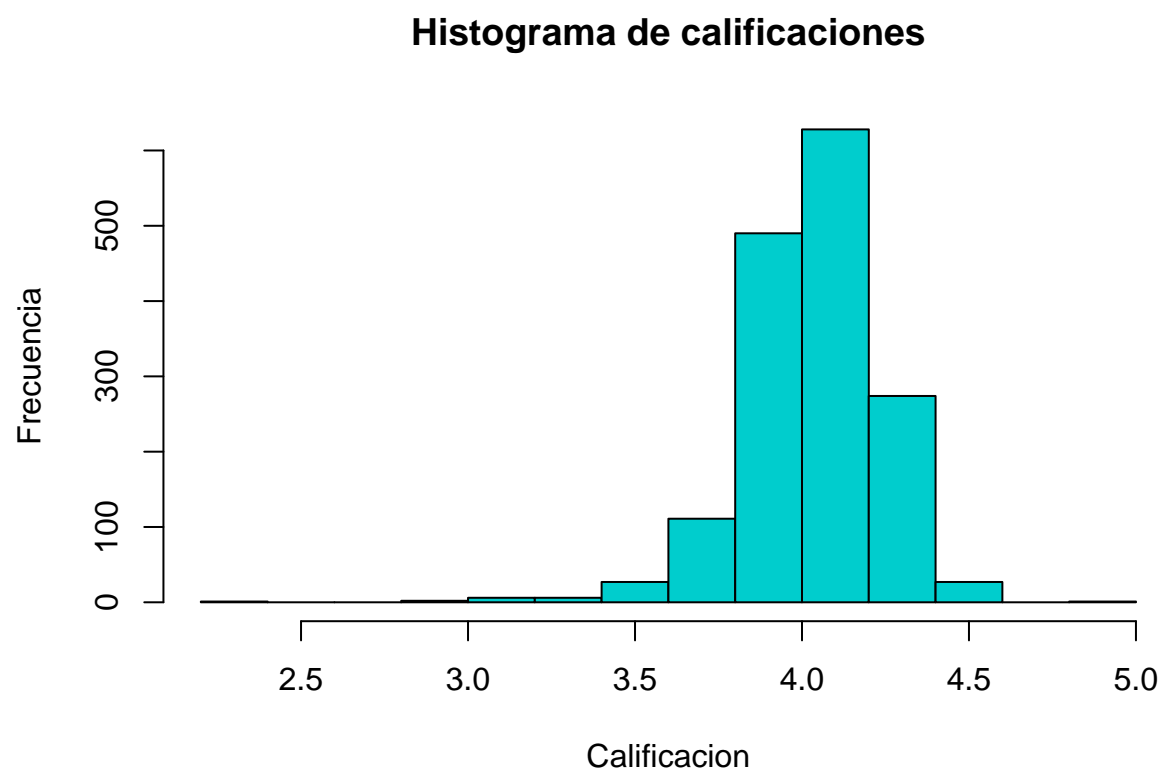


Figure 3: Histograma basico calificacion maximas de comerciantes

```
# crea un vector con las opciones de envío de los productos y realiza la frecuencia
fr = as.vector(product$precio_de_opcion_de_envio)
barplot(table(fr),
  main = "Grafico de barras envios",
  col = c("darkolivegreen1", "darkorange1", "darkorchid1", "firebrick2", "gainsboro", "darkseagreen1")
)
```



Figure 4: graficos basicos

```
pie(table(fr),
  main = "Grafico de circular envios",
  col = c("darkolivegreen1", "darkorange1", "darkorchid1", "firebrick2", "gainsboro", "darkseagreen1")
)
```

```
est = as.vector(product$tiene_banner_de_urgencia)
plot(table(fr, est),
  main = "Grafico de posaico envios",
  col = "darkolivegreen1",
  xlab = "Envio"
)
```

Grafico de circular envios

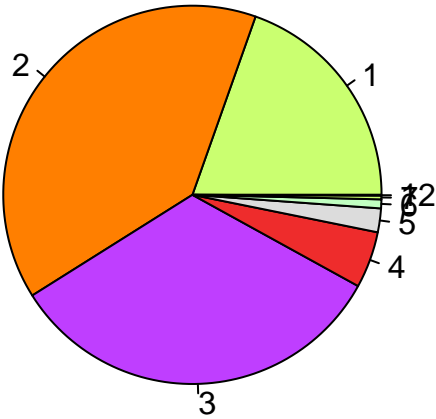


Figure 5: graficos basicos

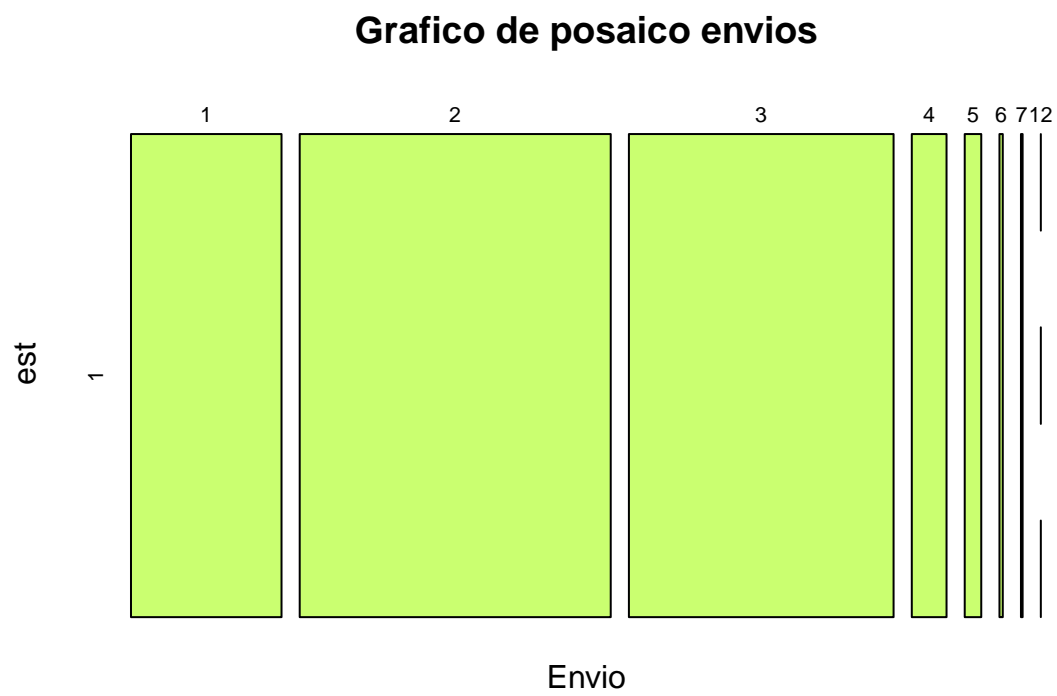


Figure 6: graficos basicos

Boxplots y outliers

```
library(dplyr)
```

```
##  
## Attaching package: 'dplyr'  
  
## The following objects are masked from 'package:stats':  
##  
##   filter, lag  
  
## The following objects are masked from 'package:base':  
##  
##   intersect, setdiff, setequal, union
```

```
library(ggplot2)  
data = read.csv("../DS/productos-verano-calificación-rendimiento_2020-08.csv")  
  
bien <- complete.cases(data)  
data <- data[bien,]  
data <- mutate(data,  
  ID_de_tamano_de_variacion_de_producto = factor(ID_de_tamano_de_variacion_de_producto),  
  paises_enviados_a = factor(paises_enviados_a))  
  
ggplot(data,  
  aes(x = ID_de_tamano_de_variacion_de_producto,  
      y = unidades_vendidas,  
      fill = paises_enviados_a)) + geom_boxplot() +  
  ggtitle("Boxplots") +  
  xlab("Categorias") +  
  ylab("Mediciones")
```

