

stio-web.R

r3172621

2025-06-02

```
# CARGAR LOS DATOS
setwd("/cloud/project/proyecto")
datos <- read.csv("archivo depurado nuevo 12.csv",
                  header = TRUE, dec = ".", sep=";")

# Variable OPERADOR
#-----

# Agrupar los valores menos frecuentes en "OTROS"
operador <- datos$OPERADOR
tabla_cruda_operador <- sort(table(operador), decreasing = TRUE)

# Mostrar los 3 valores más frecuentes, el resto se agrupa en "OTROS"
top_n_operador <- 3
top_n_valores_operador <- names(tabla_cruda_operador)[1:top_n_operador]

operador_agrupado <- ifelse(operador %in% top_n_valores_operador, operador, "OTROS")

# Nueva tabla agrupada
tabla_agrupada_operador <- table(operador_agrupado)

#-----
# TABLA DE FRECUENCIAS
#-----
Tabla_Operador <- data.frame(
  OPERADOR = names(tabla_agrupada_operador),
  ni = as.numeric(tabla_agrupada_operador)
)

Tabla_Operador$hi_porcentaje <- round((Tabla_Operador$ni / sum(Tabla_Operador$ni)) * 100, 2)

# Agregar fila TOTAL
Tabla_Operador <- rbind(
  Tabla_Operador,
  data.frame(OPERADOR = "TOTAL",
             ni = sum(Tabla_Operador$ni),
             hi_porcentaje = 100)
)

print(Tabla_Operador)
```

```
##          OPERADOR    ni hi_porcentaje
## 1          OTROS  4170          14.10
```

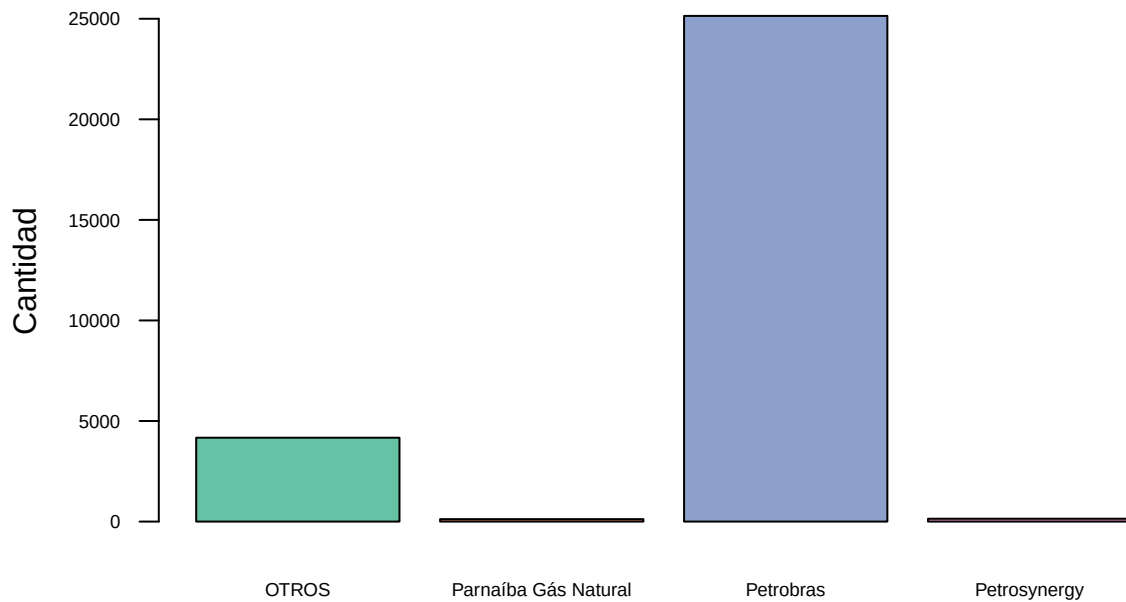
```
## 2 Parnaíba Gás Natural 121 0.41
## 3 Petrobras 25142 85.01
## 4 Petrosynergy 142 0.48
## 5 TOTAL 29575 100.00
```

```
#-----
# COLORES
#-----
library(RColorBrewer)
colores_operador <- brewer.pal(n = length(tabla_agrupada_operador), name = "Set2")

#-----
# GRÁFICOS DE BARRAS
#-----

# Frecuencia absoluta
barplot(tabla_agrupada_operador,
        main = "Frecuencia Absoluta de OPERADOR",
        col = colores_operador,
        las = 1,
        cex.names = 0.6,
        cex.axis = 0.6,
        xlab = "OPERADOR", ylab = "Cantidad")
```

Frecuencia Absoluta de OPERADOR

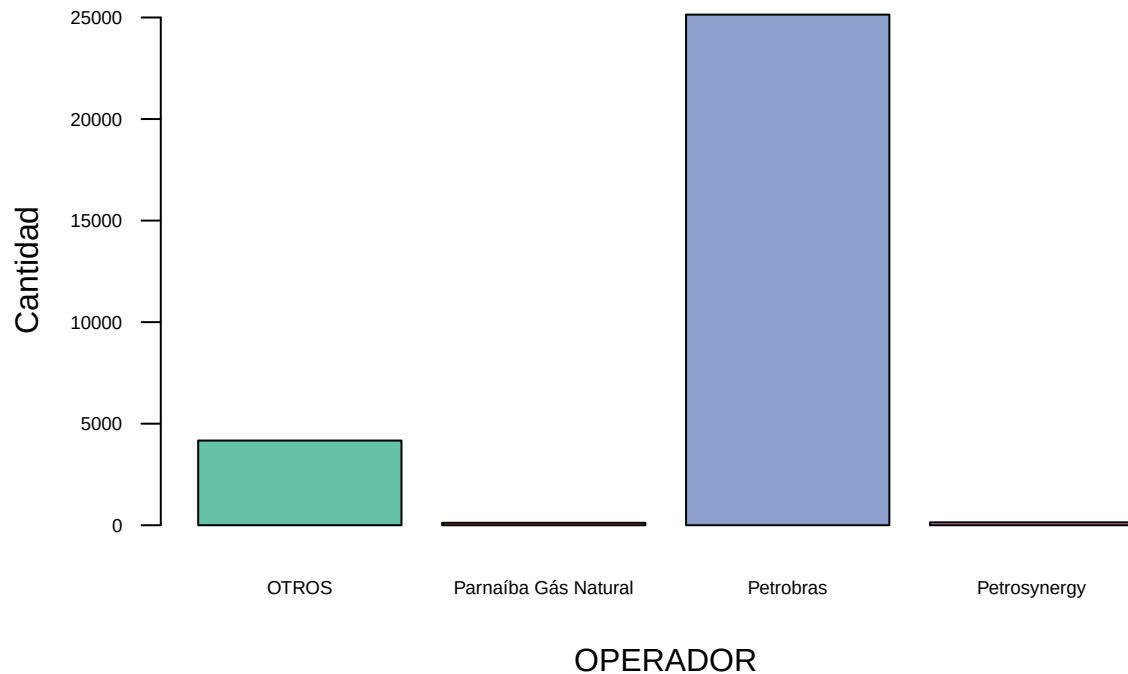


OPERADOR

```
# Frecuencia absoluta con límite Y
barplot(tabla_agrupada_operador,
        main = "Frecuencia Absoluta de OPERADOR (con límite)",
        col = colores_operador,
        las = 1,
        cex.names = 0.6,
        cex.axis = 0.6,
```

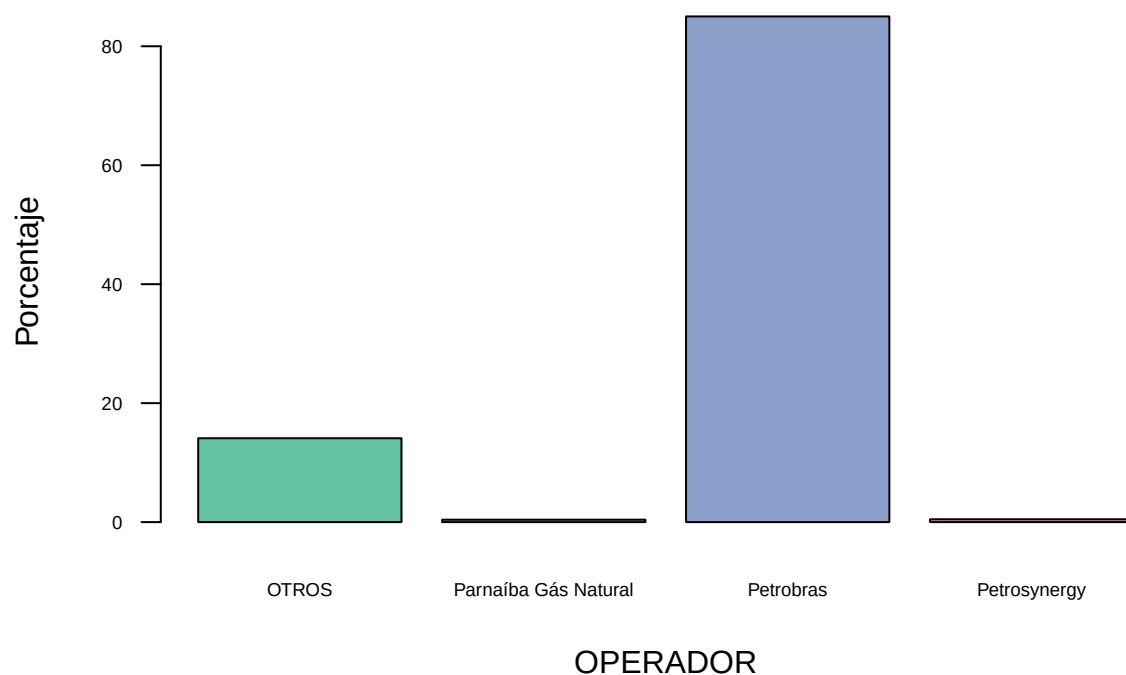
```
ylim = c(0, max(tabla_agrupada_operador) + 5),
xlab = "OPERADOR", ylab = "Cantidad")
```

Frecuencia Absoluta de OPERADOR (con límite)



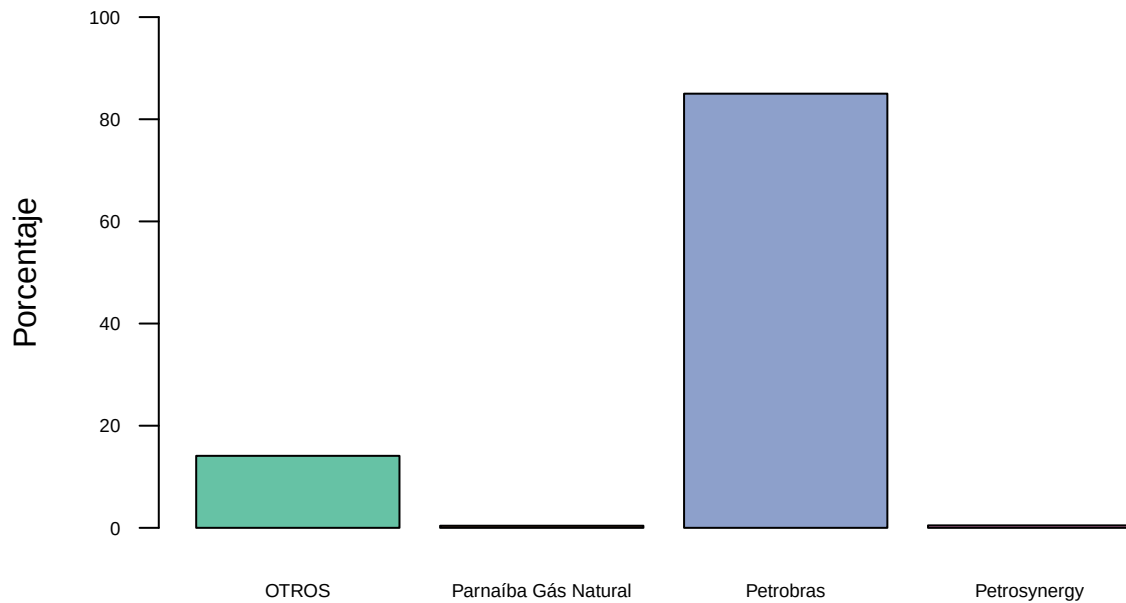
```
# Frecuencia relativa (%)
barplot(Tabla_Operador$hi_porcentaje[Tabla_Operador$OPERADOR != "TOTAL"],
main = "Frecuencia Relativa (%) de OPERADOR",
names.arg = Tabla_Operador$OPERADOR[Tabla_Operador$OPERADOR != "TOTAL"],
col = colores_operador,
las = 1,
cex.names = 0.6,
cex.axis = 0.6,
xlab = "OPERADOR", ylab = "Porcentaje")
```

Frecuencia Relativa (%) de OPERADOR



```
# Frecuencia relativa con límite Y
barplot(Tabla_Operador$hi_porcentaje[Tabla_Operador$OPERADOR != "TOTAL"],
  main = "Frecuencia Relativa (%) de OPERADOR (con límite)",
  names.arg = Tabla_Operador$OPERADOR[Tabla_Operador$OPERADOR != "TOTAL"],
  col = colores_operador,
  las = 1,
  cex.names = 0.6,
  cex.axis = 0.6,
  ylim = c(0, 100),
  xlab = "OPERADOR", ylab = "Porcentaje")
```

Frecuencia Relativa (%) de OPERADOR (con límite)



OPERADOR

```
#-----
# DIAGRAMA CIRCULAR
#-----

pie_data_operador <- Tabla_Operador$ni[Tabla_Operador$OPERADOR != "TOTAL"]
pie_labels_operador <- Tabla_Operador$OPERADOR[Tabla_Operador$OPERADOR != "TOTAL"]
pie_percent_operador <- Tabla_Operador$hi_porcentaje[Tabla_Operador$OPERADOR != "TOTAL"]

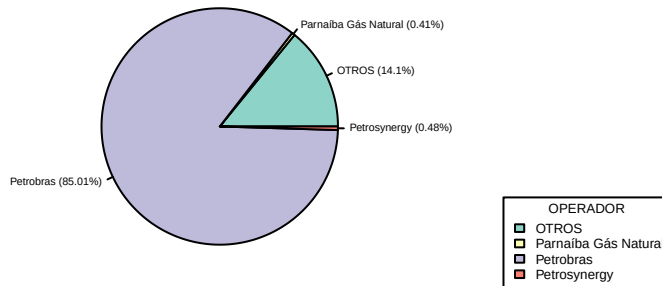
etiquetas_pie_operador <- paste0(pie_labels_operador, " (", pie_percent_operador, "%)")

n_colores_pie <- max(3, length(pie_data_operador))

pie(pie_data_operador,
  labels = etiquetas_pie_operador,
  main = "Distribución Porcentual de OPERADOR",
  col = brewer.pal(n = n_colores_pie, name = "Set3"),
  radius = 0.5,
  cex = 0.35)

legend(x = 1.2, y = -0.3,
  legend = pie_labels_operador,
  fill = brewer.pal(n = n_colores_pie, name = "Set3"),
  title = "OPERADOR",
  cex = 0.4)
```

Distribución Porcentual de OPERADOR



VARIABLE: Tipo de Ubicación

```
tipoubicacion <- datos$TERRA_MAR
TDFtipoubicacion <- table(tipoubicacion)
Tabla_Ubicacion <- as.data.frame(TDFtipoubicacion)

names(Tabla_Ubicacion) <- c("TIPO.DE.UBICACION", "ni")
Tabla_Ubicacion$hi_porcentaje <- (Tabla_Ubicacion$ni / sum(Tabla_Ubicacion$ni)) * 100

Tabla_Ubicacion <- rbind(Tabla_Ubicacion, data.frame(TIPO.DE.UBICACION = "TOTAL",
                                                    ni = sum(Tabla_Ubicacion$ni),
                                                    hi_porcentaje = 100))

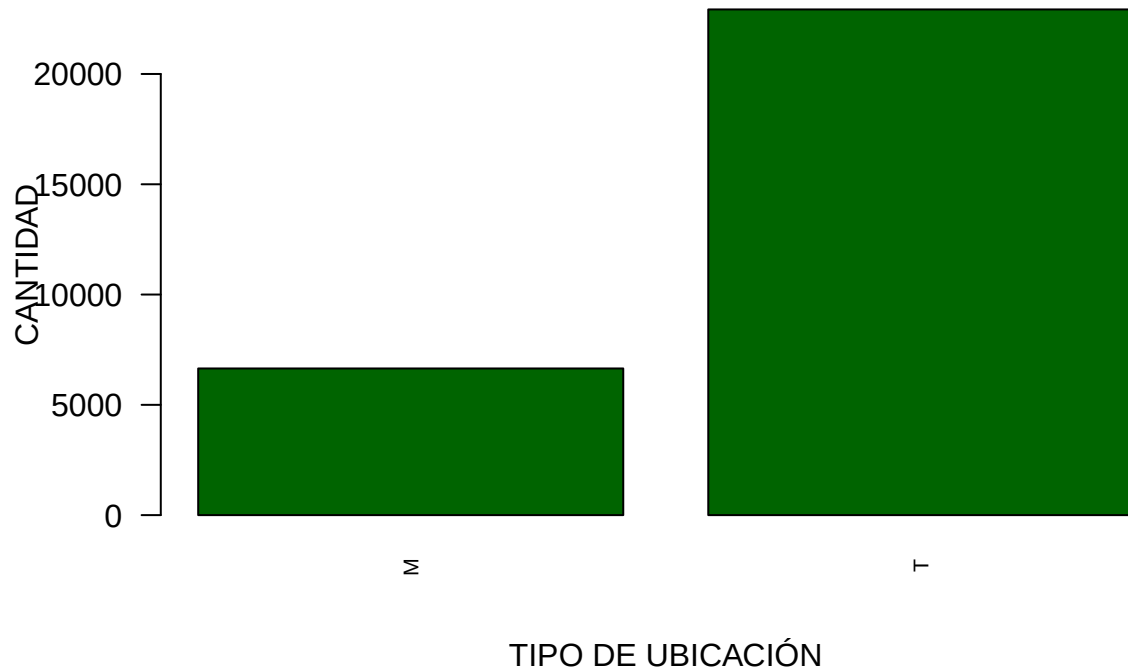
Tabla_Ubicacion
```

```
## TIPO.DE.UBICACION ni hi_porcentaje
## 1 M 6649 22.48183
## 2 T 22926 77.51817
## 3 TOTAL 29575 100.00000
```

GRÁFICAS

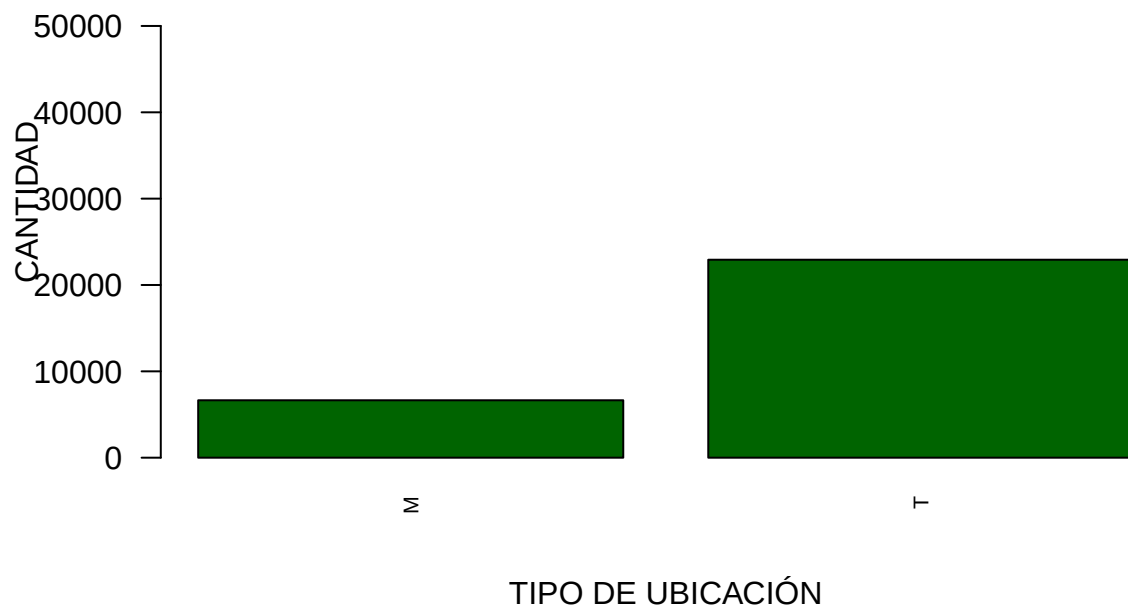
```
barplot(TDFtipoubicacion,
        main = "GRÁFICA 2.1: DISTRIBUCIÓN DE TIPO DE UBICACIÓN",
        xlab = "TIPO DE UBICACIÓN", ylab = "CANTIDAD",
        col = "darkgreen", las = 2, cex.names = 0.7)
```

GRÁFICA 2.1: DISTRIBUCIÓN DE TIPO DE UBICACIÓN



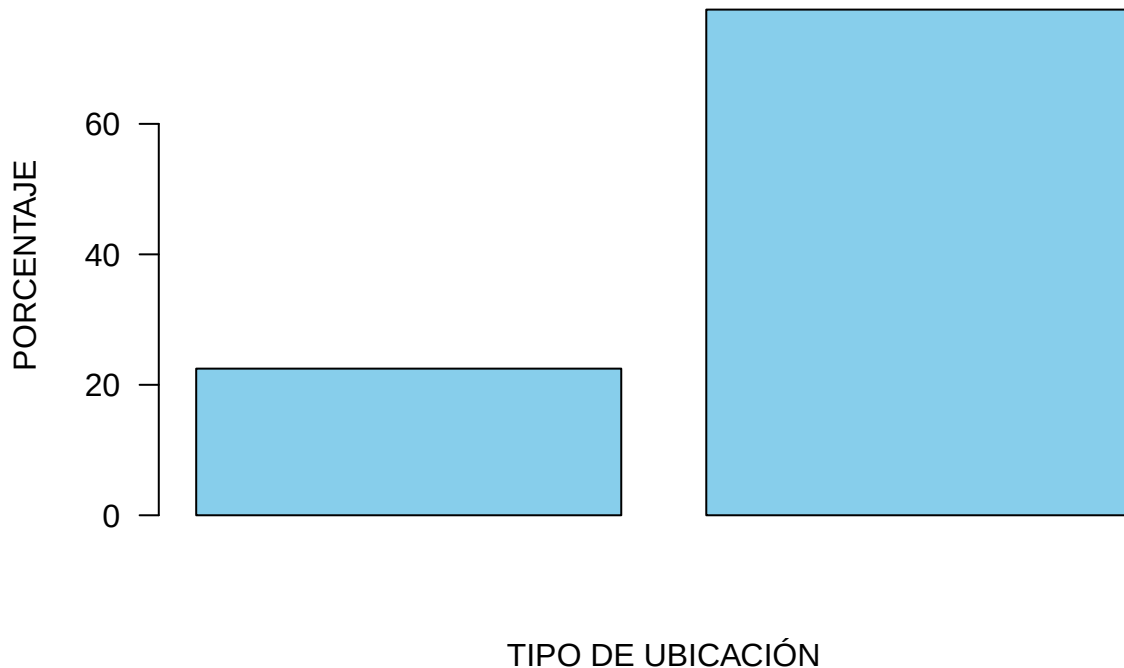
```
barplot(TDFtipoubicacion,  
  main = "GRÁFICA 2.2: DISTRIBUCIÓN DE TIPO DE UBICACIÓN",  
  xlab = "TIPO DE UBICACIÓN", ylab = "CANTIDAD",  
  col = "darkgreen", las = 2, ylim = c(0, sum(Tabla_Ubicacion$ni)), cex.names = 0.7)
```

GRÁFICA 2.2: DISTRIBUCIÓN DE TIPO DE UBICACIÓN



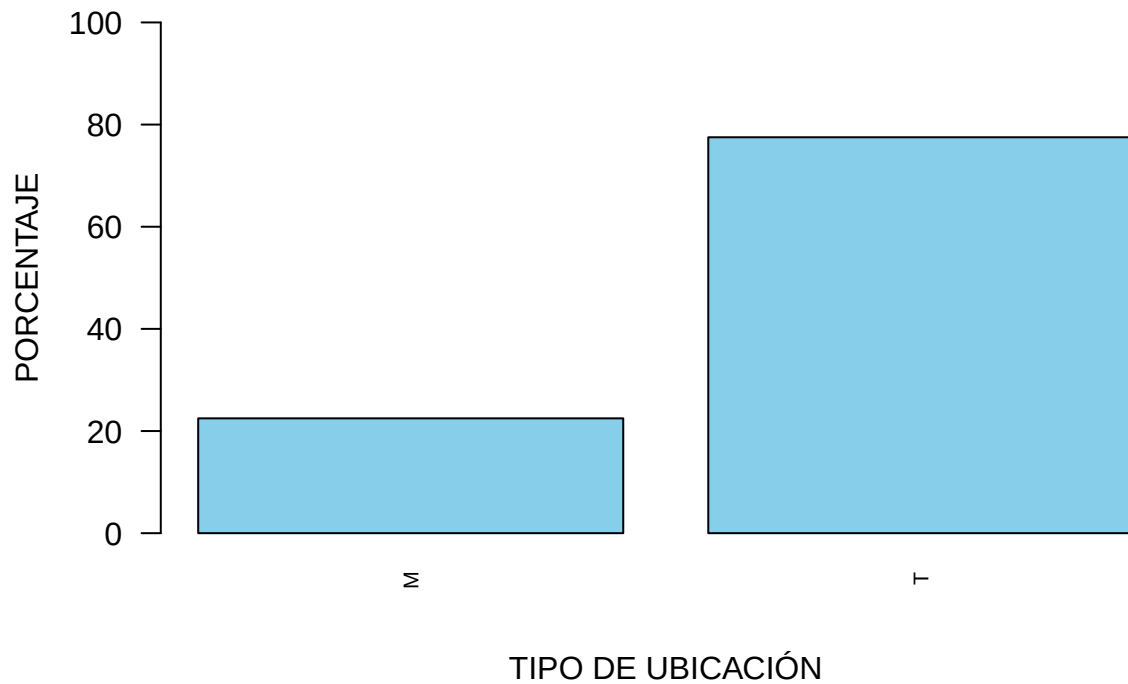
```
barplot(Tabla_Ubicacion$hi_porcentaje[Tabla_Ubicacion$TIPO.DE.UBICACION != "TOTAL"],
        main = "GRÁFICA 2.3: DISTRIBUCIÓN DE TIPO DE UBICACIÓN",
        xlab = "TIPO DE UBICACIÓN", ylab = "PORCENTAJE",
        col = "skyblue", las = 2, cex.names = 0.7)
```

GRÁFICA 2.3: DISTRIBUCIÓN DE TIPO DE UBICACIÓN



```
barplot(Tabla_Ubicacion$hi_porcentaje[Tabla_Ubicacion$TIPO.DE.UBICACION != "TOTAL"],
        main = "GRÁFICA 2.4: DISTRIBUCIÓN DE TIPO DE UBICACIÓN",
        xlab = "TIPO DE UBICACIÓN", ylab = "PORCENTAJE",
        col = "skyblue", las = 2, ylim = c(0, 100), cex.names = 0.7,
        names.arg = Tabla_Ubicacion$TIPO.DE.UBICACION[Tabla_Ubicacion$TIPO.DE.UBICACION != "TOTAL"])
```


GRÁFICA 2.4: DISTRIBUCIÓN DE TIPO DE UBICACIÓN



#Diagrama circular

```
names(Tabla_Ubicacion) <- c("TIPO.DE.UBICACION", "ni")
Tabla_Ubicacion$hi_porcentaje <- (Tabla_Ubicacion$ni / sum(Tabla_Ubicacion$ni)) * 100
```

```
pie_data_ubicacion <- Tabla_Ubicacion$ni[Tabla_Ubicacion$TIPO.DE.UBICACION != "TOTAL"]
pie_labels_ubicacion_raw <- Tabla_Ubicacion$TIPO.DE.UBICACION[Tabla_Ubicacion$TIPO.DE.UBICACION != "TOTAL"]
pie_percentages_ubicacion <- round(Tabla_Ubicacion$hi_porcentaje[Tabla_Ubicacion$TIPO.DE.UBICACION != "TOTAL"], 1)
```

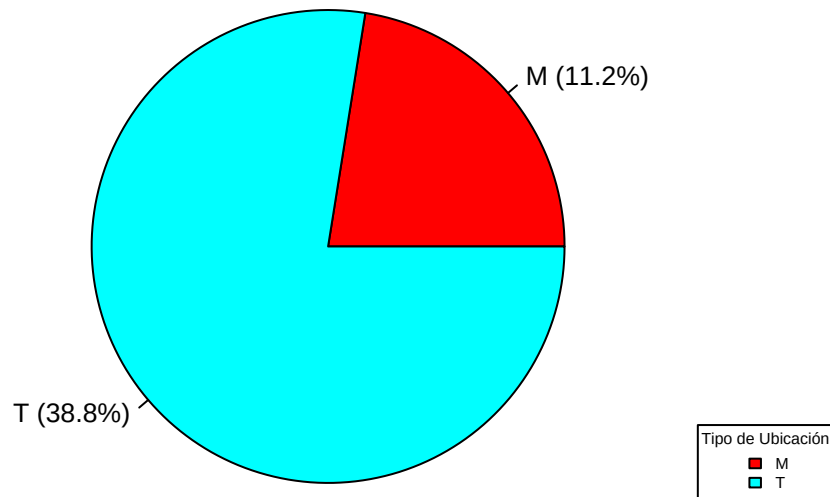
```
colores_pie_ubicacion <- rainbow(length(pie_data_ubicacion))
```

```
etiquetas_pastel_ubicacion <- paste0(pie_labels_ubicacion_raw,
                                     " (", pie_percentages_ubicacion, "%)")
```

```
pie(pie_data_ubicacion,
    labels = etiquetas_pastel_ubicacion,
    main = "Gráfica No. 2.5: Distribución Porcentual de Tipo de Ubicación",
    col = colores_pie_ubicacion,
    radius = 1,
    cex = 0.8
)
```

```
legend("bottomright",
      legend = pie_labels_ubicacion_raw,
      fill = colores_pie_ubicacion,
      title = "Tipo de Ubicación",
      cex = 0.5
)
```

Gráfica No. 2.5: Distribución Porcentual de Tipo de Ubicación



```
# VARIABLE: Tipo
tipo <- datos$TIPO
TDFtipo <- table(tipo)
Tabla_Tipo <- as.data.frame(TDFtipo)

names(Tabla_Tipo) <- c("TIPO", "ni")
Tabla_Tipo$hi_porcentaje <- (Tabla_Tipo$ni / sum(Tabla_Tipo$ni)) * 100

Tabla_Tipo <- rbind(Tabla_Tipo, data.frame(TIPO = "TOTAL",
                                           ni = sum(Tabla_Tipo$ni),
                                           hi_porcentaje = 100))

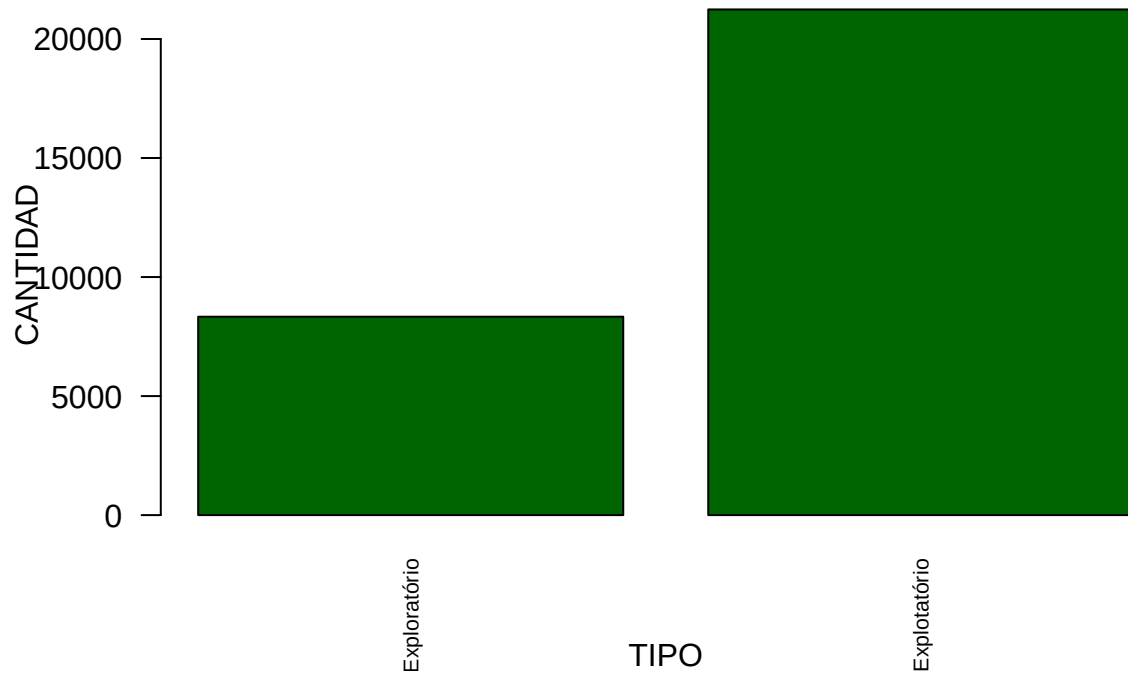
Tabla_Tipo

##          TIPO      ni hi_porcentaje
## 1 Exploratório  8334      28.17921
## 2 Explotatório 21241      71.82079
## 3          TOTAL 29575     100.00000
```

```
# GRÁFICAS para Tipo

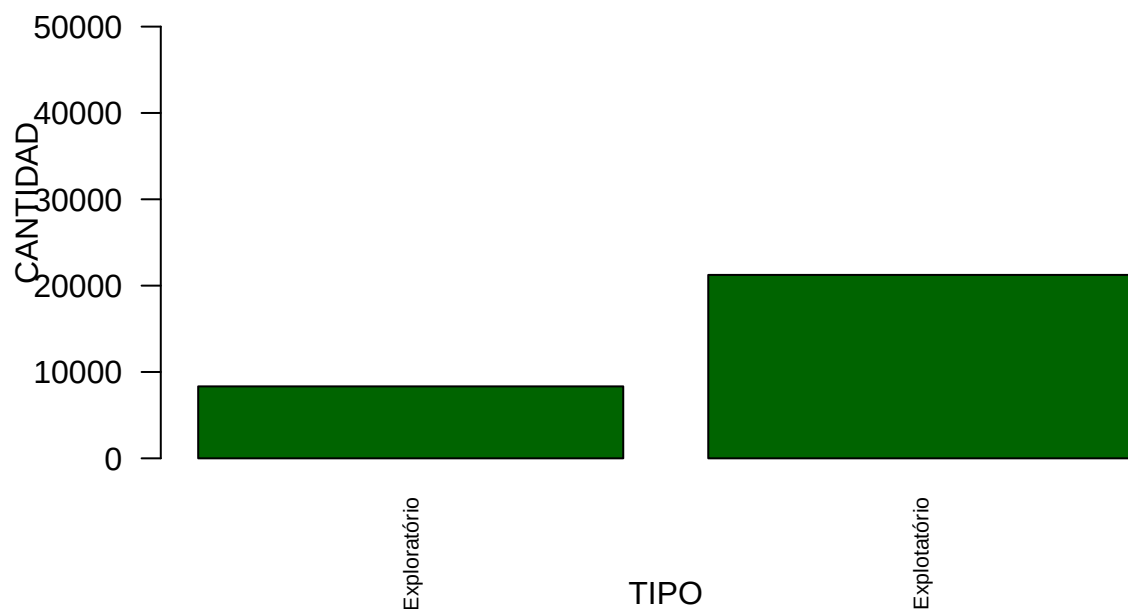
barplot(TDFtipo,
        main = "GRÁFICA 4.1: DISTRIBUCIÓN DE TIPO",
        xlab = "TIPO", ylab = "CANTIDAD",
        col = "darkgreen", las = 2, cex.names = 0.7)
```

GRÁFICA 4.1: DISTRIBUCIÓN DE TIPO



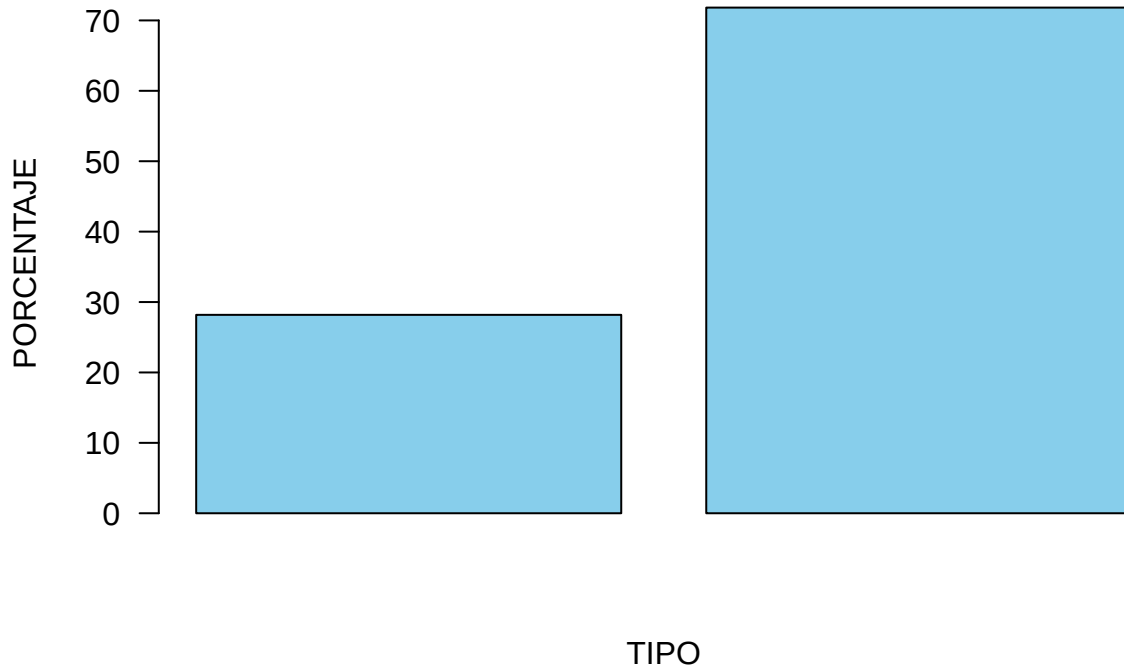
```
barplot(TDFtipo,  
  main = "GRÁFICA 4.2: DISTRIBUCIÓN DE TIPO",  
  xlab = "TIPO", ylab = "CANTIDAD",  
  col = "darkgreen", las = 2, ylim = c(0, sum(Tabla_Tipo$ni)), cex.names = 0.7)
```

GRÁFICA 4.2: DISTRIBUCIÓN DE TIPO



```
barplot(Tabla_Tipo$hi_porcentaje[Tabla_Tipo$TIPO != "TOTAL"],
        main = "GRÁFICA 4.3: DISTRIBUCIÓN DE TIPO",
        xlab = "TIPO", ylab = "PORCENTAJE",
        col = "skyblue", las = 2, cex.names = 0.7)
```

GRÁFICA 4.3: DISTRIBUCIÓN DE TIPO



```
barplot(Tabla_Tipo$hi_porcentaje[Tabla_Tipo$TIPO != "TOTAL"],
        main = "GRÁFICA 4.4: DISTRIBUCIÓN DE TIPO",
        xlab = "TIPO", ylab = "PORCENTAJE",
        col = "skyblue", las = 2, ylim = c(0, 100), cex.names = 0.7,
        names.arg = Tabla_Tipo$TIPO[Tabla_Tipo$TIPO != "TOTAL"])
```

GRÁFICA 4.4: DISTRIBUCIÓN DE TIPO

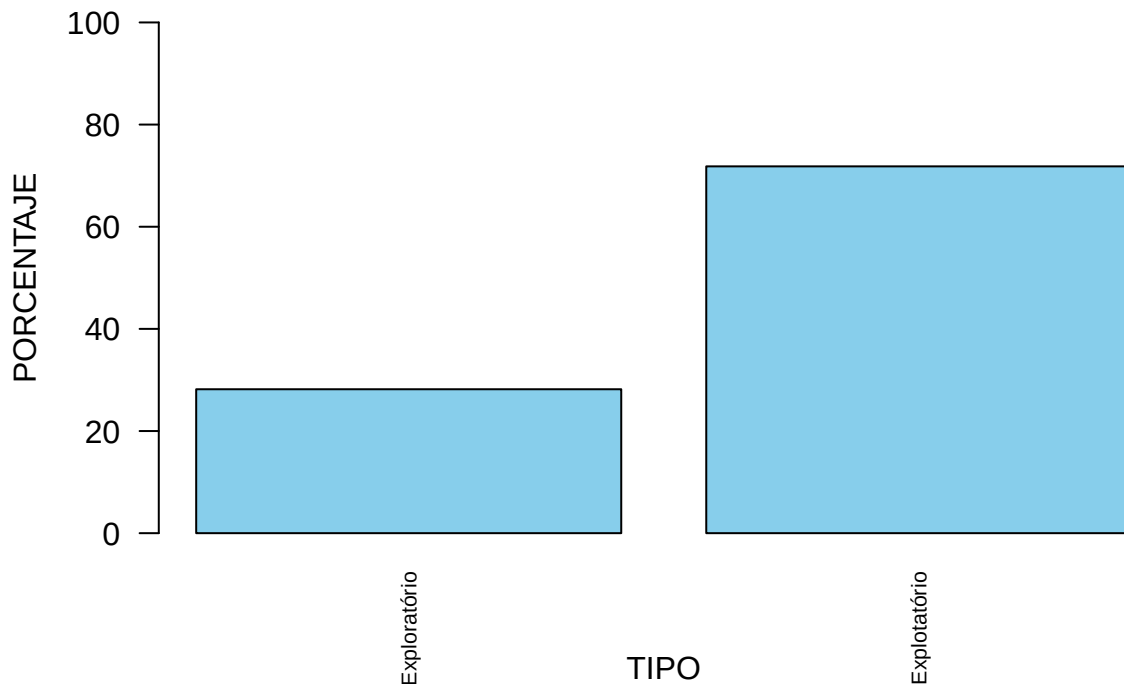


DIAGRAMA CIRCULAR para Tipo

```
pie_data_tipo <- Tabla_Tipo$ni[Tabla_Tipo$TIPO != "TOTAL"]
pie_labels_tipo_raw <- Tabla_Tipo$TIPO[Tabla_Tipo$TIPO != "TOTAL"]
pie_percentages_tipo <- round(Tabla_Tipo$hi_porcentaje[Tabla_Tipo$TIPO != "TOTAL"], 1)

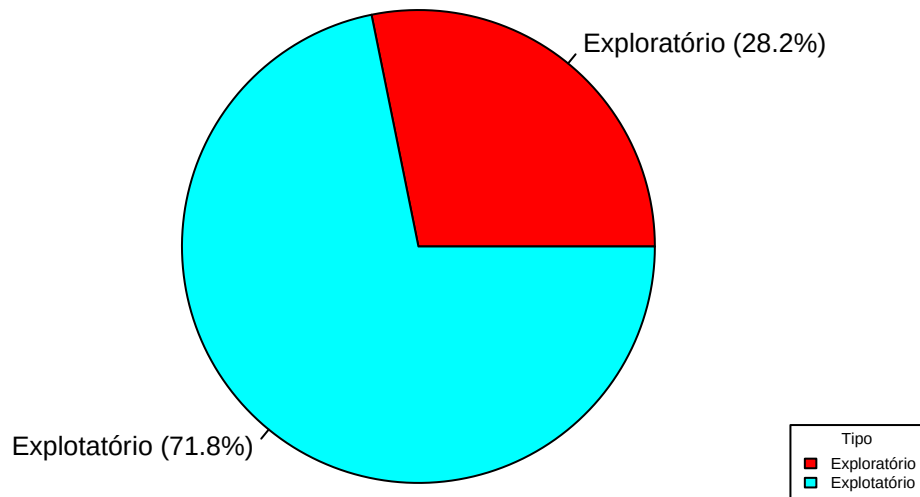
colores_pie_tipo <- rainbow(length(pie_data_tipo))

etiquetas_pastel_tipo <- paste0(pie_labels_tipo_raw,
                                " (", pie_percentages_tipo, "%)")

pie(pie_data_tipo,
    labels = etiquetas_pastel_tipo,
    main = "Gráfica No. 4.5: Distribución Porcentual de Tipo",
    col = colores_pie_tipo,
    radius = 1,
    cex = 0.8
)

legend("bottomright",
    legend = pie_labels_tipo_raw,
    fill = colores_pie_tipo,
    title = "Tipo",
    cex = 0.5
)
```

Gráfica No. 4.5: Distribución Porcentual de Tipo



```
#VARIABLE DIRECCION #####
c_direccion<-datos$DIRECAO
TDFc_direccion <- table(c_direccion)
TDFc_direccion

## c_direccion
## Direccional Horizontal Vertical
##      5560      1232      22783

#TABLA N 28
Tablac_direccion <- as.data.frame(TDFc_direccion)
hi <- Tablac_direccion$Freq/sum(Tablac_direccion$Freq)
hi

## [1] 0.1879966 0.0416568 0.7703466
sum(hi)

## [1] 1
hi_porc <- hi * 100
sum(hi_porc)

## [1] 100
Tablac_direccion <- data.frame(Tablac_direccion, hi_porc)

totaldireccion <- data.frame(
  c_direccion = "Total",
  Freq = sum(Tablac_direccion$Freq),
  hi_porc = "100%"
)

Tablac_direccion <- rbind(Tablac_direccion, totaldireccion)

colnames(Tablac_direccion)[colnames(Tablac_direccion) == "c_direccion"] <- "direccion"
```

```

names(Tablac_direccion)

## [1] "direccion" "Freq"      "hi_porc"

colnames(Tablac_direccion)[colnames(Tablac_direccion) == "Freq"] <- "ni"
names(Tablac_direccion)

## [1] "direccion" "ni"      "hi_porc"

colnames(Tablac_direccion)[colnames(Tablac_direccion) == "hi_porc"] <- "hi"
names(Tablac_direccion)

## [1] "direccion" "ni"      "hi"

#DISTRUBUCION
c_direccion <- datos$DIRECAO

TDFc_direccion_original <- table(c_direccion)
print("Tabla de Frecuencias Original para Dirección:")

## [1] "Tabla de Frecuencias Original para Dirección:"

print(TDFc_direccion_original)

## c_direccion
## Direccional Horizontal    Vertical
##      5560      1232      22783

Tablac_direccion_df <- as.data.frame(TDFc_direccion_original)
colnames(Tablac_direccion_df) <- c("direccion", "frecuencia")
Tablac_direccion_ordenada <- Tablac_direccion_df[order(Tablac_direccion_df$frecuencia, decreasing = TRUE)]

num_top_direcciones <- 12
top_direcciones <- head(Tablac_direccion_ordenada, num_top_direcciones)
frecuencia_otros_direcciones <- sum(Tablac_direccion_ordenada$frecuencia) - sum(top_direcciones$frecuencia)
otros_row_direcciones <- data.frame(direccion = "Otros", frecuencia = frecuencia_otros_direcciones)

Tablac_direccion_agrupada <- rbind(top_direcciones, otros_row_direcciones)

Tablac_direccion_agrupada$hi <- Tablac_direccion_agrupada$frecuencia / sum(Tablac_direccion_agrupada$frecuencia)
Tablac_direccion_agrupada$hi_porc <- round(Tablac_direccion_agrupada$hi * 100, 2)

print("Tabla de Frecuencias Agrupadas para Dirección (Global):")

## [1] "Tabla de Frecuencias Agrupadas para Dirección (Global):"

print(Tablac_direccion_agrupada)

##      direccion frecuencia      hi hi_porc
## 3      Vertical      22783 0.7703466   77.03
## 1    Direccional      5560 0.1879966   18.80
## 2    Horizontal      1232 0.0416568    4.17
## 11         Otros         0 0.0000000    0.00

#GRAFICAS LOCALES
#GRAFICA 01
if (length(TDFc_direccion_original) <= 25) { # Puedes ajustar este umbral si lo deseas
  bp_local_abs_direccion <- barplot(TDFc_direccion_original,
    main = "Frecuencia Absoluta de Direcciones (Local)",

```

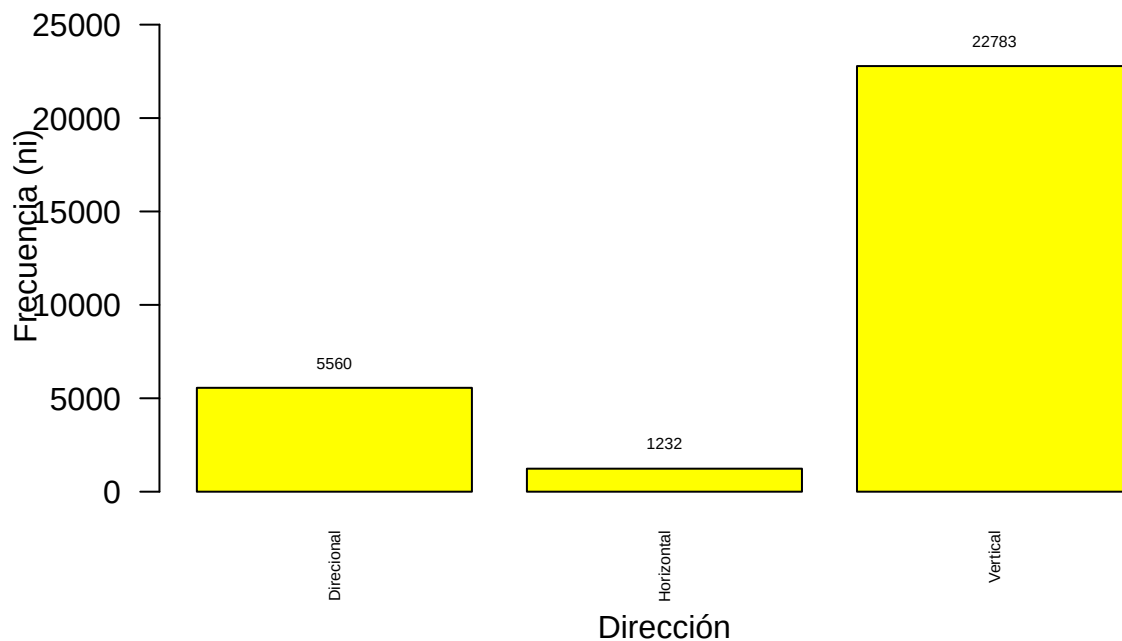
```

        xlab = "Dirección",
        ylab = "Frecuencia (ni)",
        col = "YELLOW",
        las = 2,
        cex.names = 0.5,
        ylim = c(0, max(TDFc_direccion_original) * 1.2))

text(x = bp_local_abs_direccion,
     y = TDFc_direccion_original,
     labels = TDFc_direccion_original,
     pos = 3, # Posición: encima de la barra
     cex = 0.5, # Tamaño de la etiqueta
     col = "black") # Color del texto
} else {
  print("Demasiadas categorías para un gráfico de barras 'Local' de Frecuencias Absolutas legible. Se or
}

```

Frecuencia Absoluta de Direcciones (Local)



```

#GRAFICA 02
if (length(TDFc_direccion_original) <= 25) { # Mismo umbral que el anterior
  hi_local_direccion <- TDFc_direccion_original / sum(TDFc_direccion_original)
  hi_porc_local_direccion <- round(hi_local_direccion * 100, 2)

  bp_local_porc_direccion <- barplot(hi_porc_local_direccion,
    main = "Porcentaje de Direcciones (Local)",
    xlab = "Dirección",
    ylab = "Porcentaje (%)",
    col = "BLUE",
    las = 2,
    cex.names = 0.7,
    ylim = c(0, max(hi_porc_local_direccion) * 1.2))
}

```

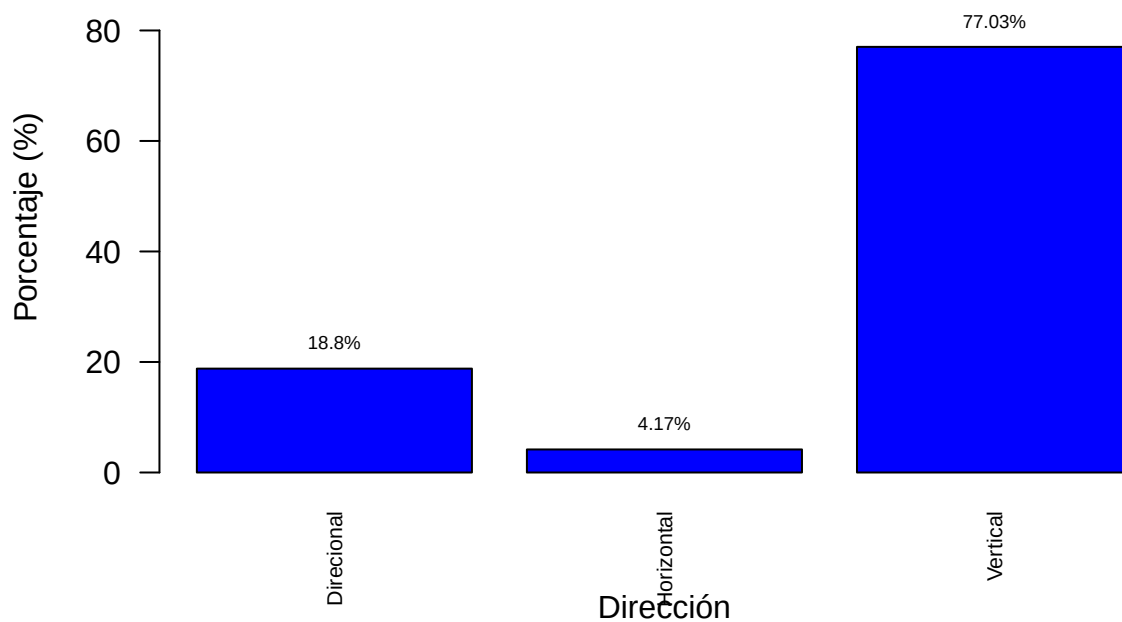


```

text(x = bp_local_porc_direccion,
     y = hi_porc_local_direccion,
     labels = paste0(hi_porc_local_direccion, "%"),
     pos = 3,
     cex = 0.6,
     col = "black")
} else {
  print("Demasiadas categorías para un gráfico de barras 'Local' de Porcentajes legible. Se omite.")
}

```

Porcentaje de Direcciones (Local)



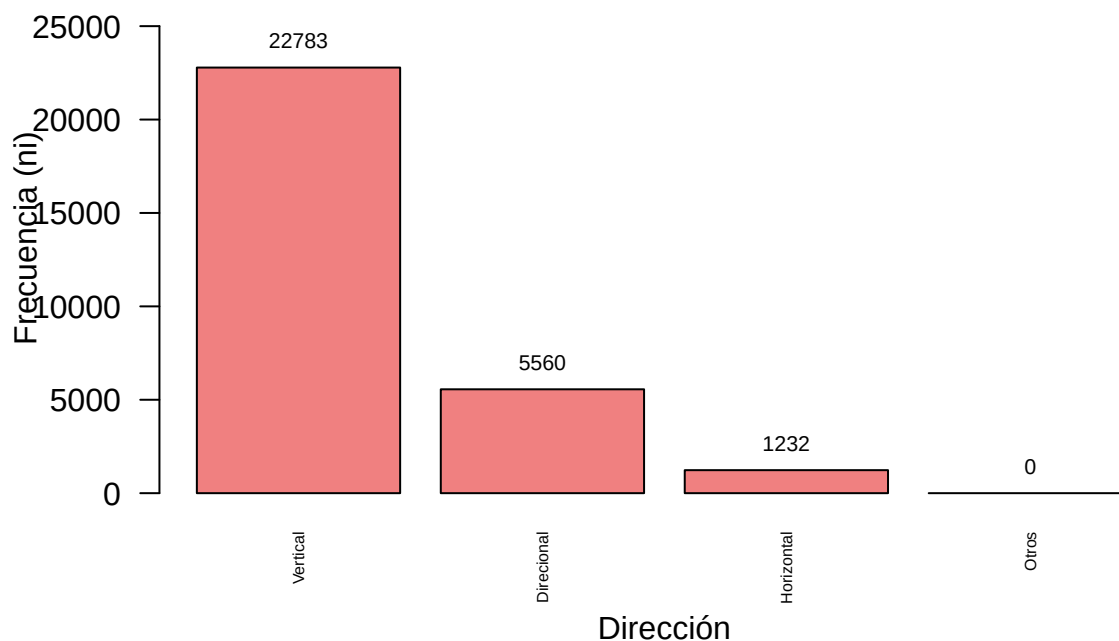
```

#GRAFICAS GLOBALES
#GRAFICA 01
bp_global_abs_direccion <- barplot(Tabla direccion_agrupada$frecuencia,
                                   names.arg = Tabla direccion_agrupada$direccion,
                                   main = "Frecuencia Absoluta de Direcciones (Global - Agrupado)",
                                   xlab = "Dirección",
                                   ylab = "Frecuencia (ni)",
                                   col = "lightcoral",
                                   las = 2,
                                   cex.names = 0.5,
                                   ylim = c(0, max(Tabla direccion_agrupada$frecuencia) * 1.2))

text(x = bp_global_abs_direccion,
     y = Tabla direccion_agrupada$frecuencia,
     labels = Tabla direccion_agrupada$frecuencia,
     pos = 3,
     cex = 0.7,
     col = "black")

```

Frecuencia Absoluta de Direcciones (Global – Agrupado)

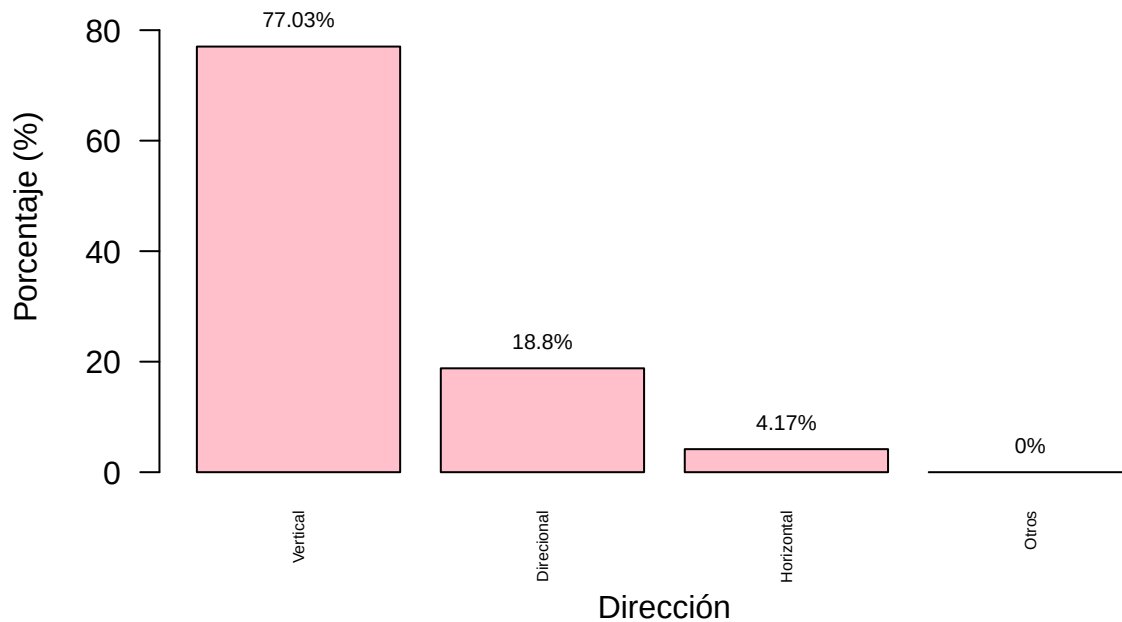


#GRAFICA 02

```
bp_global_porc_direccion <- barplot(Tablac_direccion_agrupada$hi_porc,
                                     names.arg = Tablac_direccion_agrupada$direccion,
                                     main = "Porcentaje de Direcciones (Global – Agrupado)",
                                     xlab = "Dirección",
                                     ylab = "Porcentaje (%)",
                                     col = "PINK",
                                     las = 2,
                                     cex.names = 0.5,
                                     ylim = c(0, max(Tablac_direccion_agrupada$hi_porc) * 1.2))

text(x = bp_global_porc_direccion,
     y = Tablac_direccion_agrupada$hi_porc,
     labels = paste0(Tablac_direccion_agrupada$hi_porc, "%"),
     pos = 3,
     cex = 0.7,
     col = "black")
```

Porcentaje de Direcciones (Global – Agrupado)

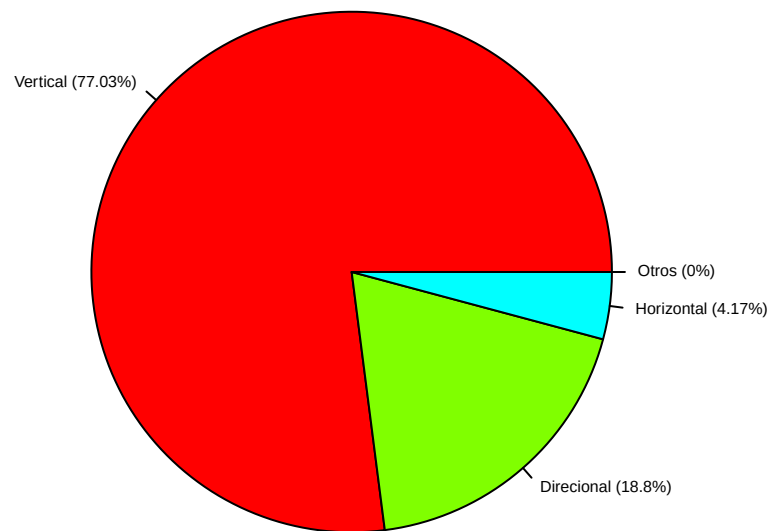


```
#DIAGRAMA CIRCULAR
pie_labels_global_direccion <- paste0(Tablac_direccion_agrupada$direccion, " (", Tablac_direccion_agrupada$frecuencia, "%)", sep="")
par_original <- par(no.readonly = TRUE)

par(mar = c(2, 2, 2, 2) + 0.1) # Puedes experimentar con estos valores, valores más bajos dan más espacio

pie(Tablac_direccion_agrupada$frecuencia,
    labels = pie_labels_global_direccion,
    main = "Distribución Porcentual de Direcciones (Global – Agrupado)",
    col = rainbow(nrow(Tablac_direccion_agrupada)),
    cex = 0.5)
```

Distribución Porcentual de Direcciones (Global – Agrupado)



```
par(par_original)
```