

Análisis del comportamiento suicida en el departamento del Tolima (2013 - 2021)

Cuaderno de análisis de datos

Carolina Ortiz López
Universidad San Buenaventura

Wednesday, December 11, 2024

Table of contents

| | |
|--------------------------------------------------------|----|
| 1 Instalación y cargar de paquetes | 1 |
| 2 Cargar los datos | 2 |
| 3 Reestructuración de los datos: | 2 |
| 4 Modelamiento estadístico | 3 |
| 4.1 Transformar los datos | 3 |
| 4.2 Ajuste del modelo | 4 |
| 4.3 Visualización de resultados | 4 |
| 5 Visualización de los resultados | 6 |
| 6 Visualización por municipio | 7 |
| 6.1 Instalar y cargar los paquetes | 7 |
| 6.2 Crear coordenadas de los municipios | 7 |
| 6.3 Visualizar mapa del Tolima | 8 |
| 6.4 Convertir las coordenadas del dataset | 10 |
| 6.5 Visualizar el número de suicidios e intentos | 11 |
| 6.6 Mapa interactivo | 13 |

1 Instalación y cargar de paquetes

```
if (!require("brms")) install.packages("brms")
if (!require("dplyr")) install.packages("dplyr")
if (!require("ggalluvial")) install.packages("ggalluvial")
if (!require("ggplot2")) install.packages("ggplot2")
if (!require("glmmTMB")) install.packages("glmmTMB")
if (!require("here")) install.packages("here")
if (!require("htmlwidgets")) install.packages("htmlwidgets")
if (!require("leaflet")) install.packages("leaflet")
if (!require("performance")) install.packages("performance")
if (!require("readxl")) install.packages("readxl")
if (!require("tidyverse")) install.packages("tidyverse")
```

2 Cargar los datos

A continuación cargaremos los datos recopilados desde el archivo “Data/Cond_Suicida_2013_2021” que contiene información de marcadores de la Casos_Suicidio y Casos_Intento del departamento del Tolima para el periodo 2013 - 2021.

```
data <- read_xlsx (here("Data/Cond_Suicida_2013_2021.xlsx"))
head(data)
```

```
# A tibble: 6 × 10
  No. Municipio Casos_Suicidio Tasa_Mortalidad Suicidio_Hombres
  <dbl> <chr>          <dbl>          <dbl>          <dbl>
1     1 Cajamarca         30          17.5          23.7
2     2 Alpujarra          7          16.3           5.53
3     3 Santa Isabel        8          14.3           6.33
4     4 Rovira            26          13.9          20.6
5     5 Icononzo          14          13.8          11.1
6     6 Falan             10          13.4           7.91
# i 5 more variables: Suicidio_Mujeres <dbl>, Casos_Intento <dbl>,
# Tasa_Incidencia <dbl>, Intento_Hombres <dbl>, Intento_Mujeres <dbl>
```

Vamos a analizar el comportamiento suicida en hombres y mujeres. En primero lugar, debemos reestructurar la tabla sustancialmente.

3 Reestrcturación de los datos:

Vamos a seleccionar las columnas relevantes y a generar una tabla larga para hombres y mujeres teneindo en cuanto los intesntos y los suicidios completados.

```
library(tidyr)

# Reshape the data to long format for counts
data_long <- data %>%
  select(Municipio, Suicidio_Hombres, Suicidio_Mujeres, Intento_Hombres,
  Intento_Mujeres) %>%
  pivot_longer(
    cols = starts_with("Suicidio_"),
    names_to = "Sexo",
    names_prefix = "Suicidio_",
    values_to = "Completados"
  ) %>%
  left_join(
    data %>%
      select(Municipio, Intento_Hombres, Intento_Mujeres) %>%
      pivot_longer(
        cols = starts_with("Intento_"),
        names_to = "Sexo",
```

```

    names_prefix = "Intento_",
    values_to = "Intentos"
  ),
  by = c("Municipio", "Sexo")
)

str(data_long)

```

```

tibble [94 × 6] (S3: tbl_df/tbl/data.frame)
 $ Municipio      : chr [1:94] "Cajamarca" "Cajamarca" "Alpujarra" "Alpujarra" ...
 $ Intento_Hombres: num [1:94] 76.69 76.69 6.25 6.25 18.34 ...
 $ Intento_Mujeres: num [1:94] 107.31 107.31 8.75 8.75 25.66 ...
 $ Sexo           : chr [1:94] "Hombres" "Mujeres" "Hombres" "Mujeres" ...
 $ Completados    : num [1:94] 23.72 6.28 5.53 1.47 6.33 ...
 $ Intentos       : num [1:94] 76.69 107.31 6.25 8.75 18.34 ...

```

Podemos apreciar que generamos nuevas columnas denominadas “Sexo”, “Completados” e “Intentos”. Con la tabla transformada podemos realizar el modelamiento estadístico. Dejaremos la visualización de los datos (gráficos) para después. Estamos interesados en modelar la distribución de los suicidios entre hombres y mujeres, condicionando en los intentos.

4 Modelamiento estadístico

4.1 Transformar los datos

Para modelar adecuadamente estas variables, los casos e intentos de suicidios deben ser enteros (números completos, no fracciones).

```

data_long$Completados <- as.integer(data_long$Completados)
data_long$Intentos <- as.integer(data_long$Intentos)

str(data_long)

```

```

tibble [94 × 6] (S3: tbl_df/tbl/data.frame)
 $ Municipio      : chr [1:94] "Cajamarca" "Cajamarca" "Alpujarra" "Alpujarra" ...
 $ Intento_Hombres: num [1:94] 76.69 76.69 6.25 6.25 18.34 ...
 $ Intento_Mujeres: num [1:94] 107.31 107.31 8.75 8.75 25.66 ...
 $ Sexo           : chr [1:94] "Hombres" "Mujeres" "Hombres" "Mujeres" ...
 $ Completados    : int [1:94] 23 6 5 1 6 1 20 5 11 2 ...
 $ Intentos       : int [1:94] 76 107 6 8 18 25 57 81 21 29 ...

```

Podemos apreciar que “Completados” e “Intentos” son ahora enteros.

4.2 Ajuste del modelo

Para este caso ajustaremos un modelo binomial, donde los suicidios son los casos (completados, succed) de intentos de suicidio. Utilizaremos una distribución binomial con una función link “logit” para estimar las probabilidades (log-odds) de que los intentos de suicidio terminen en suicidio en función del sexo. La regresión toma la siguiente notación:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \times \text{Sexo}$$

Donde P es la probabilidad de completar el suicidio dado un intento. β_0 es la probabilidad (log-odds) de completar el suicidio para los hombres (la categoría de referencia); y β_1 representa el cambio de probabilidad (log-odds) de completar el suicidio para las mujeres en relación con los hombres.

```
Suicidios_Mdl1 <- bf(Completados | trials(Intentos) ~ Sexo)

get_prior(Suicidios_Mdl1, data_long, family = binomial)
```

```
              prior      class      coef group resp dpar nlpar lb ub
              (flat)         b              SexoMujeres
student_t(3, 0, 2.5) Intercept
      source
      default
(vectorized)
      default
```

```
Suicidios_Fit1 <-
  brm(
    data = data_long,
    family = binomial,
    formula = Suicidios_Mdl1,
    chains = 4,
    cores = 4,
    warmup = 2500,
    iter = 5000,
    seed = 8807,
    control = list(adapt_delta = 0.99, max_treedepth = 15),
    file = here("Models/Suicidio_Completados_Fit1.rds"),
    file_refit = "never"
  )
```

4.3 Visualización de resultados

A continuación visualizamos la tabla de resultados:

```
summary(Suicidios_Fit1)
```

```
Family: binomial
Links: mu = logit
Formula: Completados | trials(Intentos) ~ Sexo
Data: data_long (Number of observations: 94)
Draws: 4 chains, each with iter = 5000; warmup = 2500; thin = 1;
       total post-warmup draws = 10000
```

Regression Coefficients:

| | Estimate | Est.Error | l-95% CI | u-95% CI | Rhat | Bulk_ESS | Tail_ESS |
|-------------|----------|-----------|----------|----------|------|----------|----------|
| Intercept | -0.97 | 0.04 | -1.05 | -0.88 | 1.00 | 6511 | 5309 |
| SexoMujeres | -2.03 | 0.09 | -2.21 | -1.86 | 1.00 | 3425 | 4357 |

Draws were sampled using sampling(NUTS). For each parameter, Bulk_ESS and Tail_ESS are effective sample size measures, and Rhat is the potential scale reduction factor on split chains (at convergence, Rhat = 1).

Los resultados indican que los hombres poseen una probabilidad (log-odds) de -0.97 de completar el suicidio. Para las mujeres esta probabilidad es 2.03 veces menor. Realizaremos una estimación más directa de las probabilidades.

Dado que los hombres

$$\text{logit}(p_{\text{men}}) = -0.97$$

podemos calcular que la probabilidad de completar un suicidio es:

$$p_{\text{men}} = \frac{e^{-0.97}}{1 + e^{-0.97}} \approx \frac{0.379}{1 + 0.379} \approx 0.2747$$

Esto quiere decir que el 27.5% de los hombres en el departamento del tolima culminan sus intentos de suicidio

$$\text{Odds}_{\text{women}} = e^{\text{logit}(p_{\text{women}})} = e^{-3.00} \approx 0.0498$$

Por el contrario los log-odds para las mujeres es:

$$\text{logit}(p_{\text{women}}) = -0.97 + (-2.03) = -3.00$$

Lo que indica una probabilidad de

$$p_{\text{women}} = \frac{e^{-3.00}}{1 + e^{-3.00}} \approx \frac{0.0498}{1 + 0.0498} \approx 0.0474$$

Esto quiere decir que aproximadamente un 4% de las mujeres completan sus intentos de suicidio.

Ahora bien, si comparamos el odds ratio de las mujeres y hombres

$$OR = e^{-2.03} \approx 0.1315$$

obtenemos que las mujeres tienen aproximadamente 0.13 veces probabilidad (odds) de completar el suicidio que los hombres. Alternativamente se puede decir que ($1/0.1315 = 7.6$) los hombre tienen una probabilidad 7.6 veces mayor de completar suicidios que las mujeres.

5 Visualización de los resultados

Ahora veremos gráficamente estas estimaciones utilizando conditional effects

```
Suicidios_CE <- conditional_effects(Suicidios_Fit1)

Suicidios_CE_Fig <- plot(Suicidios_CE, plot = FALSE)[[1]]

Suicidios_CE_Fig <- Suicidios_CE_Fig +
  labs(title = "Efecto del sexo en la consumación de intentos de suicidio",
        y = "Probabilidad de suicidio dados los intentos",
        x = "Sexo")+
  theme_classic()

ggsave (Suicidios_CE_Fig ,
        file = here("Plots/Suicidio_Sexo.jpg"),
        width = 15,
        height = 10,
        units = "cm")

Suicidios_CE_Fig
```

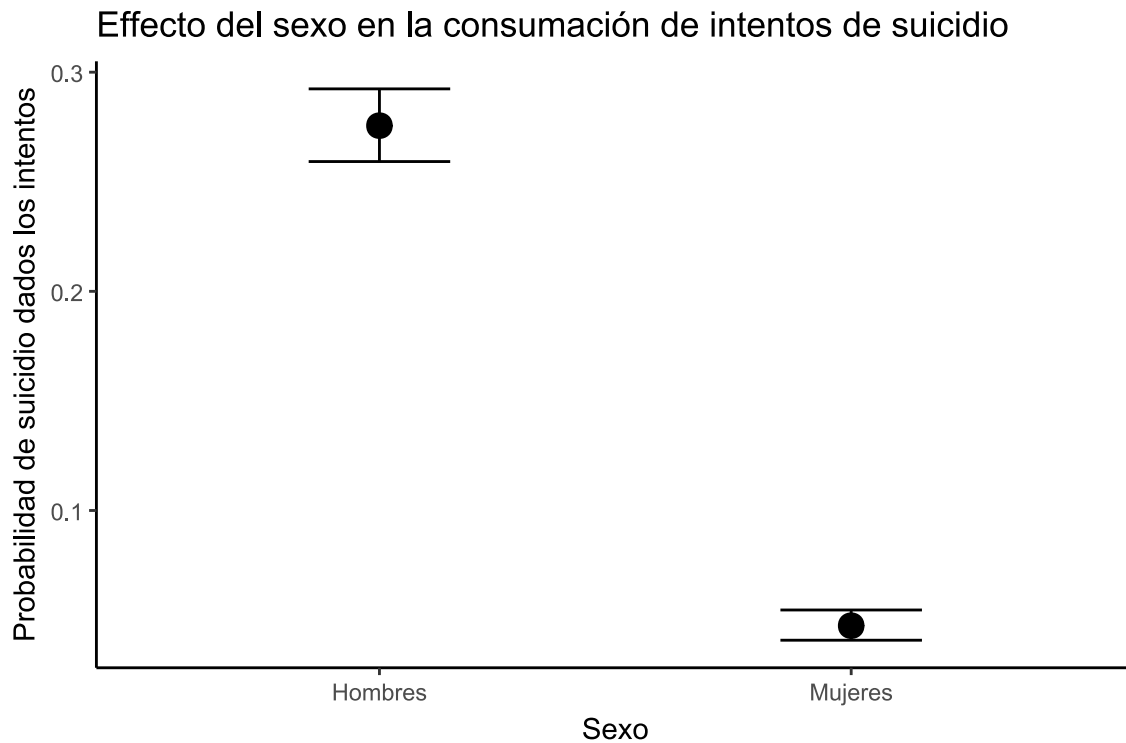


Figure 1: Efecto del sexo en la consumación de intentos de suicidio

6 Visualización por municipio

En esta sección tomaremos los resultados para realizar una visualización por municipio utilizando los paquetes “rnatuarearth” y “rnatuarearthdata”.

6.1 Instalar y cargar los paquetes

```
if (!require("rnatuarearth")) install.packages("rnatuarearth")
if (!require("rnatuarearthdata")) install.packages("rnatuarearthdata")
if (!require("sf")) install.packages("sf")
if (!require("viridis")) install.packages("viridis")

if (!require("devtools")) install.packages("devtools")
library(devtools)
devtools::install_github("ropensci/rnatuarearthhires")
```

6.2 Crear coordenadas de los municipios

Ahora, adjuntamos las coordenadas de los municipios al data frame con el que hemos estado trabajando.

```
# Creamos un data frame con las coordenadas de los municipios.
coordinates <- data.frame(
```

```

Municipio = c("Cajamarca", "Alpujarra", "Santa Isabel", "Rovira", "Icononzo",
"Falan",
              "Carmen de Apicalá", "Alvarado", "Murillo", "Villarrica", "San
Antonio",
              "Palocabildo", "Roncesvalles", "Honda", "Herveo", "Mariquita",
"Libano",
              "Melgar", "Cunday", "Fresno", "Venadillo", "Suárez", "Armero
Guayabal",
              "Ambalema", "Ibagué", "Chaparral", "Espinal", "Lérida", "Coyaima",
"Ataco",
              "Planadas", "Anzóategui", "Dolores", "Guamo", "Purificación",
"Coello",
              "Ortega", "San Luis", "Rioblanco", "Natagaima", "Valle de San
Juan", "Piedras",
              "Casabianca", "Villahermosa", "Saldaña", "Prado", "Flandes"),
Latitude = c(4.4423, 3.417, 4.713, 4.239, 4.183, 5.083, 4.073, 4.567, 4.874,
3.933,
              3.920, 5.118, 4.016, 5.208, 5.084, 5.198, 4.921, 4.204, 4.061,
5.152,
              4.690, 4.047, 5.031, 4.782, 4.4389, 3.723, 4.150, 4.861, 3.833,
3.590,
              3.200, 4.630, 3.533, 4.030, 3.858, 4.283, 3.938, 5.203, 3.521,
3.622,
              4.288, 4.498, 5.081, 5.030, 3.930, 3.755, 4.289),
Longitude = c(-75.4287, -74.917, -75.097, -75.247, -74.533, -74.960, -74.720,
-74.950,
              -75.170, -74.600, -75.467, -75.029, -75.607, -74.735, -75.170,
-74.891,
              -75.063, -74.640, -74.693, -75.035, -74.934, -74.831, -74.891,
-74.764,
              -75.2322, -75.483, -74.880, -74.910, -75.083, -75.501, -75.643,
-75.102,
              -74.733, -74.970, -74.933, -74.933, -75.220, -74.923, -75.463,
-75.094,
              -75.114, -74.875, -75.124, -75.120, -75.020, -74.927, -74.817)
)

# Juntamos los dos datasets
data_long <- left_join(data_long, coordinates, by = "Municipio")

```

6.3 Visualizar mapa del Tolima

Ahora utilizamos los paquetes de geolocalización para ver el mapa del departamento del Tolima con los municipios del del data frame.

```

library(ggplot2)
library(rnaturalearth)
library(rnaturalearthdata)

```



```

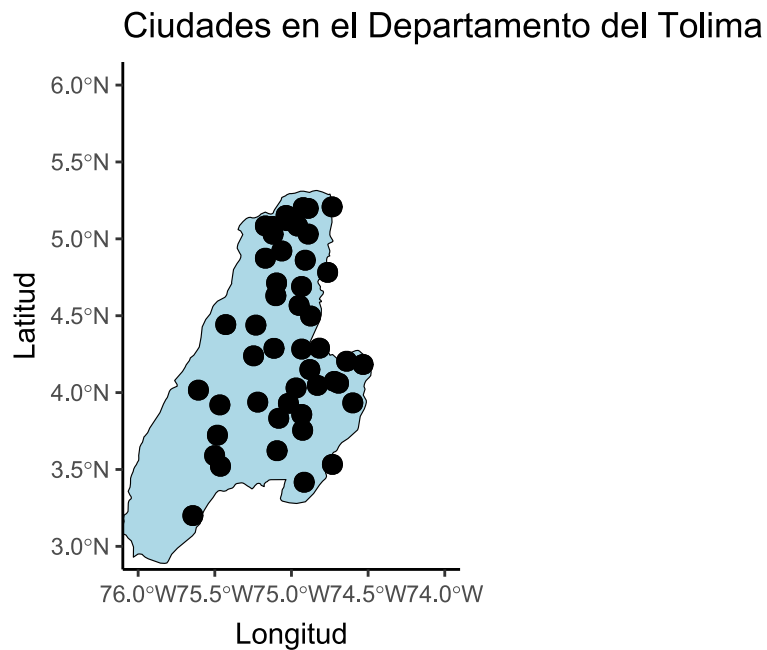
library(sf)
library(dplyr)
library(readxl)

# Descargamos los mapas de Colombia
colombia_map <- ne_countries(scale = "medium", country = "Colombia", returnclass = "sf")
col_departments <- ne_states(country = "Colombia", returnclass = "sf")

# Filtramos en la base de datos de mapas al departamento del tolima
tolima_map <- col_departments %>% filter(name == "Tolima")

# Creamos el mapa de las ciudades
ggplot(data = tolima_map) +
  geom_sf(fill = "lightblue", color = "black") +
  geom_point(data = data_long, aes(x = Longitude, y = Latitude), color = "black",
size = 3) + # Plot cities
  labs(title = "Ciudades en el Departamento del Tolima, Colombia", x = "Longitud",
y = "Latitud") +
  coord_sf(xlim = c(-76, -74), ylim = c(3, 6)) +
  theme_classic()

```



Ahora, verificamos que nuestros datos y el mapa del tolima usan las mismas coordenadas de referencia (CRS). Esto es crucial para las operaciones espaciales que realizaremos.

```
st_crs(tolima_map)
```

```

Coordinate Reference System:
  User input: WGS 84
  wkt:
GEOGCRS["WGS 84",
  DATUM["World Geodetic System 1984",
    ELLIPSOID["WGS 84",6378137,298.257223563,
      LENGTHUNIT["metre",1]],
    PRIMEM["Greenwich",0,
      ANGLEUNIT["degree",0.0174532925199433]],
  CS[ellipsoidal,2],
    AXIS["latitude",north,
      ORDER[1],
      ANGLEUNIT["degree",0.0174532925199433]],
    AXIS["longitude",east,
      ORDER[2],
      ANGLEUNIT["degree",0.0174532925199433]],
  ID["EPSG",4326]]

```

6.4 Convertir las coodenadas del dataset

Ahora, convertimos nuestroso datos en un objeto `sf`, especificando las coordenadas y el CRS.

```

data_long_sf <- st_as_sf(
  data_long,
  coords = c("Longitude", "Latitude"),
  crs = 4326, # Mantemos lel sistema de referencia WGS84
  remove = FALSE
)

```

Con el siguiente código, nos aseguramos que nuestros datos y el mapa tengan las mismas coordenadas:

```

st_crs(data_long_sf)

```

```

Coordinate Reference System:
  User input: EPSG:4326
  wkt:
GEOGCRS["WGS 84",
  ENSEMBLE["World Geodetic System 1984 ensemble",
    MEMBER["World Geodetic System 1984 (Transit)"],
    MEMBER["World Geodetic System 1984 (G730)"],
    MEMBER["World Geodetic System 1984 (G873)"],
    MEMBER["World Geodetic System 1984 (G1150)"],
    MEMBER["World Geodetic System 1984 (G1674)"],
    MEMBER["World Geodetic System 1984 (G1762)"],
    MEMBER["World Geodetic System 1984 (G2139)"],

```

```

    ELLIPSOID["WGS 84",6378137,298.257223563,
      LENGTHUNIT["metre",1]],
    ENSEMBLEACCURACY[2.0]],
    PRIMEM["Greenwich",0,
      ANGLEUNIT["degree",0.0174532925199433]],
    CS[ellipsoidal,2],
      AXIS["geodetic latitude (Lat)",north,
        ORDER[1],
        ANGLEUNIT["degree",0.0174532925199433]],
      AXIS["geodetic longitude (Lon)",east,
        ORDER[2],
        ANGLEUNIT["degree",0.0174532925199433]],
    USAGE[
      SCOPE["Horizontal component of 3D system."],
      AREA["World."],
      BBOX[-90,-180,90,180]],
    ID["EPSG",4326]]

```

```

# Si las coordenadas son diferentes, se homogenizan
if (st_crs(data_long_sf) != st_crs(tolima_map)) {
  data_long_sf <- st_transform(data_long_sf, crs = st_crs(tolima_map))
}

```

6.5 Visualizar el número de suicidios e intentos

Ahora, utilizamos ggplot para visualizar el número de suicidios por municipio en el Departamento del Tolima.

```

# Creamos el mapa base de Tolima
base_map <- ggplot() +
  geom_sf(data = tolima_map, fill = "gray90", color = "black")

# Adjuntamos puntos como probabilidades
Suicide_map <- base_map +
  geom_sf(
    data = data_long_sf,
    aes(geometry = geometry,
      color = Intentos,
      size = Completados),

  ) +
  scale_color_viridis_c(option = "plasma",
    limits = c(0, 200)) +

  scale_size(limits = c(0, 50)) +

  theme_minimal() +

```

```

theme(legend.position = "bottom") +

labs(
  title = "Suicidios consumados en el Departamento del Tolima",
  color = "Intentos"
) +

geom_text(
  data = data_long_sf,
  aes(x = Longitude, y = Latitude, label = Municipio),
  size = 4, color = "black", nudge_y = 0.05
) +

facet_wrap(~ Sexo)

ggsave (Suicide_map,
  file = here("Plots/Suicidio_Ciudad.jpg"),
  width = 24,
  height = 20,
  units = "cm")

Suicide_map

```

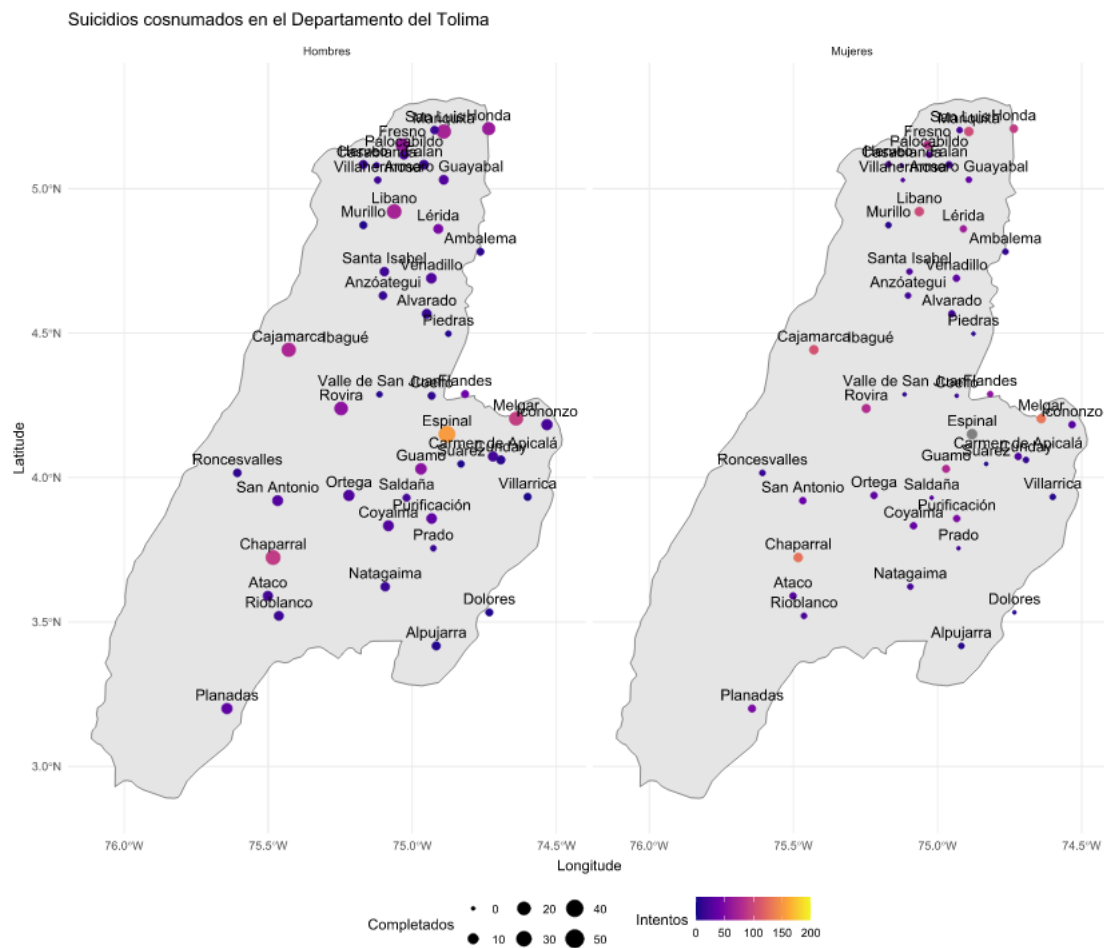


Figure 2: Suicidios de hombres y mujeres en el Departamento del Tolima

Figure 2 nos muestra el numero de suicidios consumados con el tamaño de los puntos y los intentos con código de color.

6.6 Mapa interactivo

Podemos también utilizar el paquete `leaflet` para hacer un mapa interactivo. En primer lugar, creamos un mapa centrado en el departamento del Tolima.

```
map <- leaflet() %>%
  addTiles() %>% # Add default OpenStreetMap tiles
  setView(lng = mean(data_long_sf$Longitude), lat = mean(data_long_sf$Latitude),
    zoom = 8)
```

A continuación, utilizamos el mapa para agregar capas y visualizar la distribución de intentos y suicidios en cada uno de los municipios del departamento del Tolima.

```

# Create color palette
pal <- colorNumeric(
  palette = "Plasma",
  domain = data_long_sf$Intentos
)

# Separate data by Sexo
data_hombres <- data_long_sf %>% filter(Sexo == "Hombres")
data_mujeres <- data_long_sf %>% filter(Sexo == "Mujeres")

# Adicionamos las capas
Suicidio_Mapa <- leaflet() %>%
  addTiles() %>%
  setView(lng = mean(data_long_sf$Longitude), lat = mean(data_long_sf$Latitude),
    zoom = 8) %>%
  addPolygons(
    data = tolima_map,
    fillColor = "transparent",
    color = "blue",
    weight = 2,
    opacity = 1,
    group = "Tolima Boundary"
  ) %>%
  addCircleMarkers(
    data = data_hombres,
    lng = ~Longitude,
    lat = ~Latitude,
    radius = ~sqrt(Completados) * 2,
    stroke = TRUE,
    fillOpacity = 0.7,
    popup = ~paste0(
      "<strong>Municipio: </strong>", Municipio, "<br>",
      "<strong>Sexo: </strong>", Sexo, "<br>",
      "<strong>Intentos: </strong>", Intentos, "<br>",
      "<strong>Completados: </strong>", Completados, "<br>"
    ),
    group = "Hombres"
  ) %>%
  addCircleMarkers(
    data = data_mujeres,
    lng = ~Longitude,
    lat = ~Latitude,
    radius = ~sqrt(Completados) * 2,
    stroke = TRUE,
    fillOpacity = 0.7,
    popup = ~paste0(
      "<strong>Municipio: </strong>", Municipio, "<br>",
      "<strong>Sexo: </strong>", Sexo, "<br>",

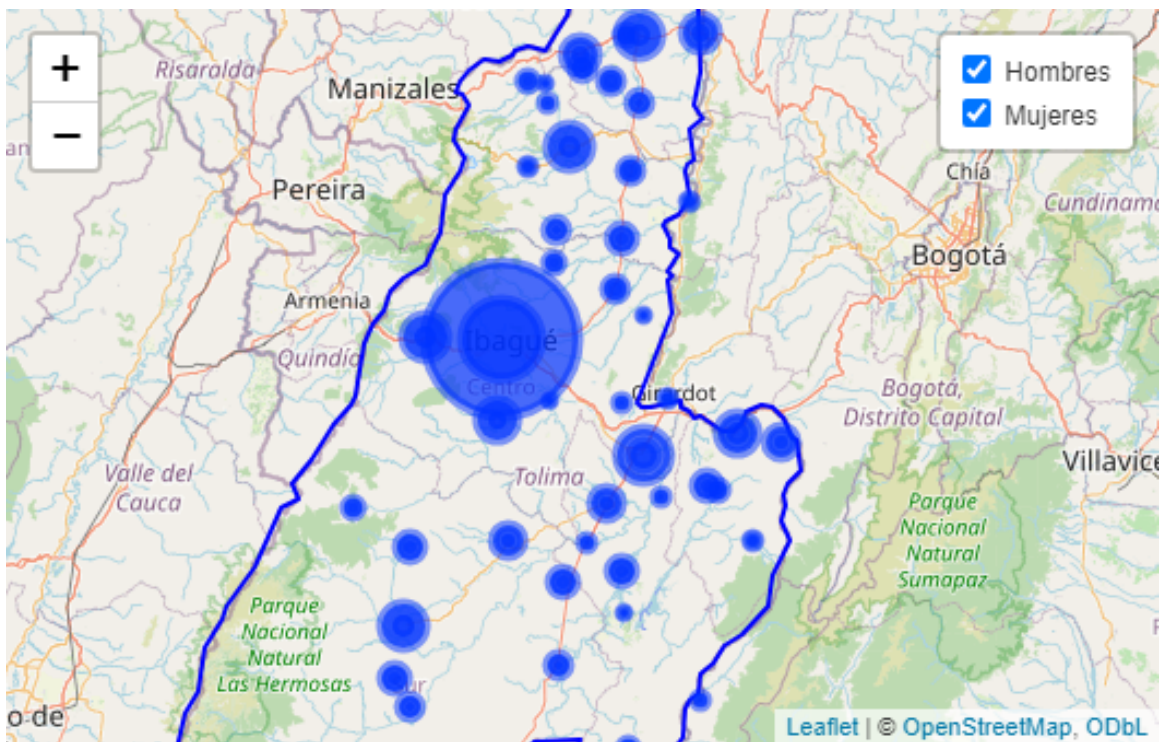
```

```

    "<strong>Intentos: </strong>", Intentos, "<br>",
    "<strong>Completados: </strong>", Completados, "<br>"
  ),
  group = "Mujeres"
) %>%
addLayersControl(
  overlayGroups = c("Hombres", "Mujeres"),
  options = layersControlOptions(collapsed = FALSE)
)

```

Suicidio_Mapa



El mapa nos permite ver la información en cada municipio separado por hombres y mujeres.