

# Visualizaciones Preliminares: Negociado de Policía de Puerto Rico

Félix A Báez Santiago

2024-01-23

## Imports

```
library(tidyverse)
library(readxl)
library(kableExtra)
library(zoo)
library(here)
path <- here("data", "Negociado_de_Policia", "/")
```

Funcion para manejar datos missing

## Cantidad de mujeres desaparecidas, localizadas y sin localizar- adultas y menores (desde enero 2020 a marzo 2023)

```
## Joining, by = c("Categoria", "Enero", "Febrero", "Marzo", "Abril", "Mayo", "Junio",
## "Julio", "Agosto", "Septiembre", "Octubre", "Noviembre", "Diciembre", "Total",
## "Año")
## Joining, by = c("Categoria", "Enero", "Febrero", "Marzo", "Abril", "Mayo", "Junio",
## "Julio", "Agosto", "Septiembre", "Octubre", "Noviembre", "Diciembre", "Total",
## "Año")

## # A tibble: 216 x 6
##   Categoria      Año Meses      Casos Casos_Normalizados Fecha
##   <chr>          <chr> <fct>      <dbl>          <dbl> <yearmon>
## 1 Adultas Desaparecidas 2020 Enero         13             2 Jan 2020
## 2 Adultas Desaparecidas 2020 Febrero        18          2.77 Feb 2020
## 3 Adultas Desaparecidas 2020 Marzo         15          2.31 Mar 2020
## 4 Adultas Desaparecidas 2020 Abril          8          1.23 Apr 2020
## 5 Adultas Desaparecidas 2020 Mayo          16          2.46 May 2020
## 6 Adultas Desaparecidas 2020 Junio          12          1.85 Jun 2020
## 7 Adultas Desaparecidas 2020 Julio           14          2.15 Jul 2020
## 8 Adultas Desaparecidas 2020 Agosto          16          2.46 Aug 2020
## 9 Adultas Desaparecidas 2020 Septiembre      18          2.77 Sep 2020
## 10 Adultas Desaparecidas 2020 Octubre         10          1.54 Oct 2020
## # ... with 206 more rows
```

## Resumen Estadístico

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Categoria    5  11240   2248.1   151.4 <2e-16 ***
## Residuals   210    3117    14.8
## ---
```

Table 1: Resumen estadístico por categoría

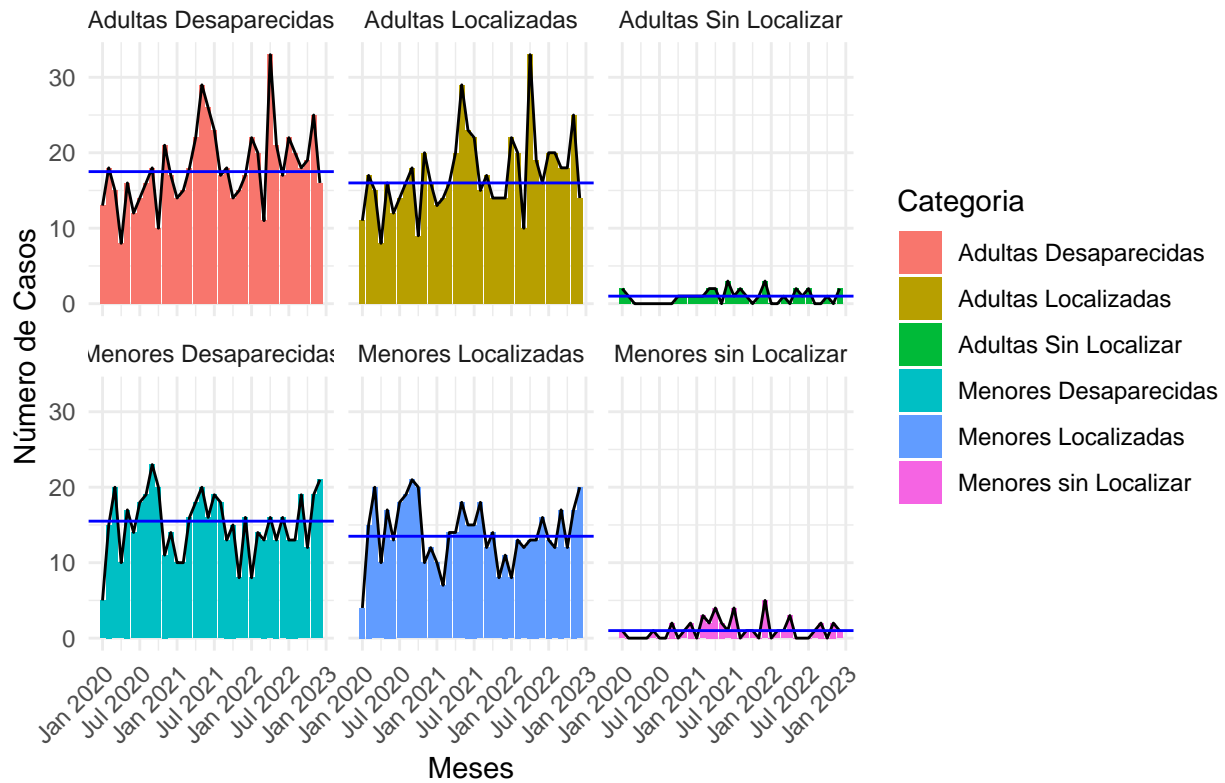
Categoría	Media	Mediana	Desviación Estándar
Adultas Desaparecidas	18.0555556	17.5	5.1377841
Adultas Localizadas	17.1666667	16.0	5.1796304
Adultas Sin Localizar	0.8888889	1.0	0.9189366
Menores Desaparecidas	15.0555556	15.5	4.1261843
Menores Localizadas	13.9166667	13.5	4.0240349
Menores sin Localizar	1.1388889	1.0	1.3342259

## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Conclusión: Los resultados de ANOVA muestran una diferencia significativa entre al menos dos de las medias de las categorías ( $F = 151$ ,  $p < 0.0000000000000002$ ), lo que sugiere que al menos una de las categorías tiene una media significativamente diferente a las demás.

## Distribución de Casos a lo largo del año

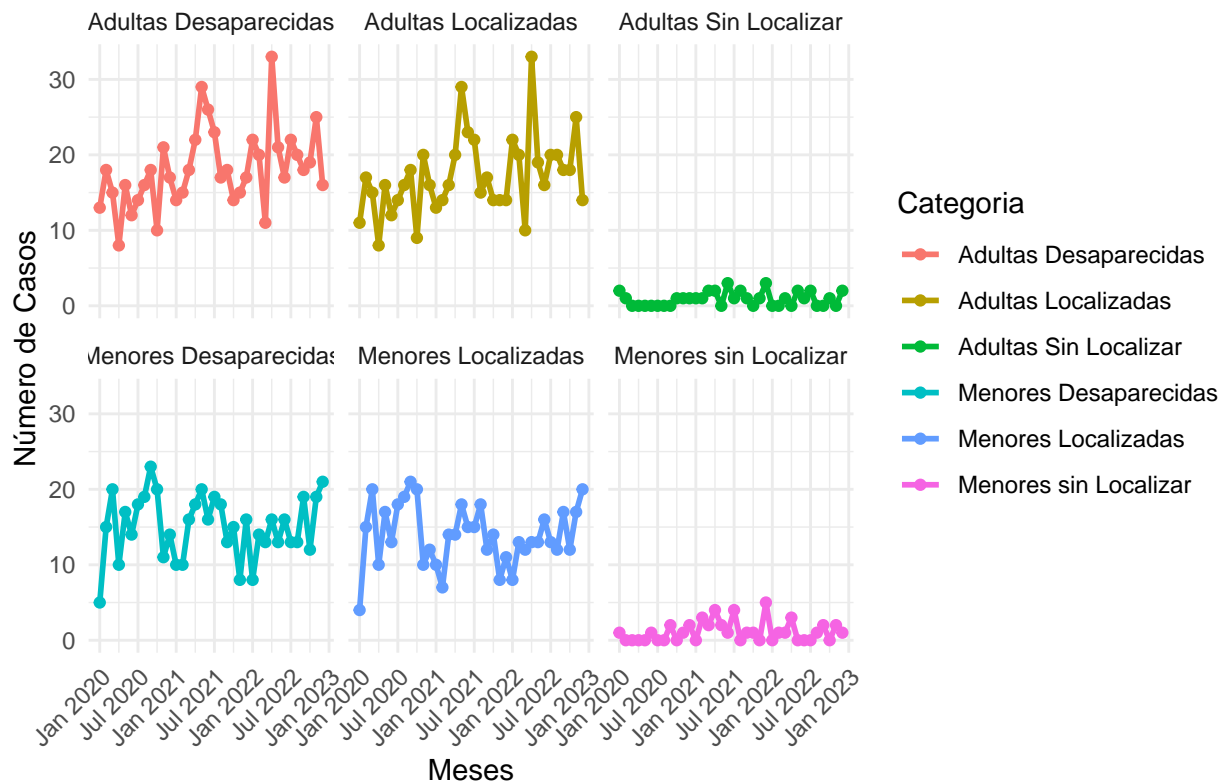
### 1: Tendencia de por Mes y Categoría



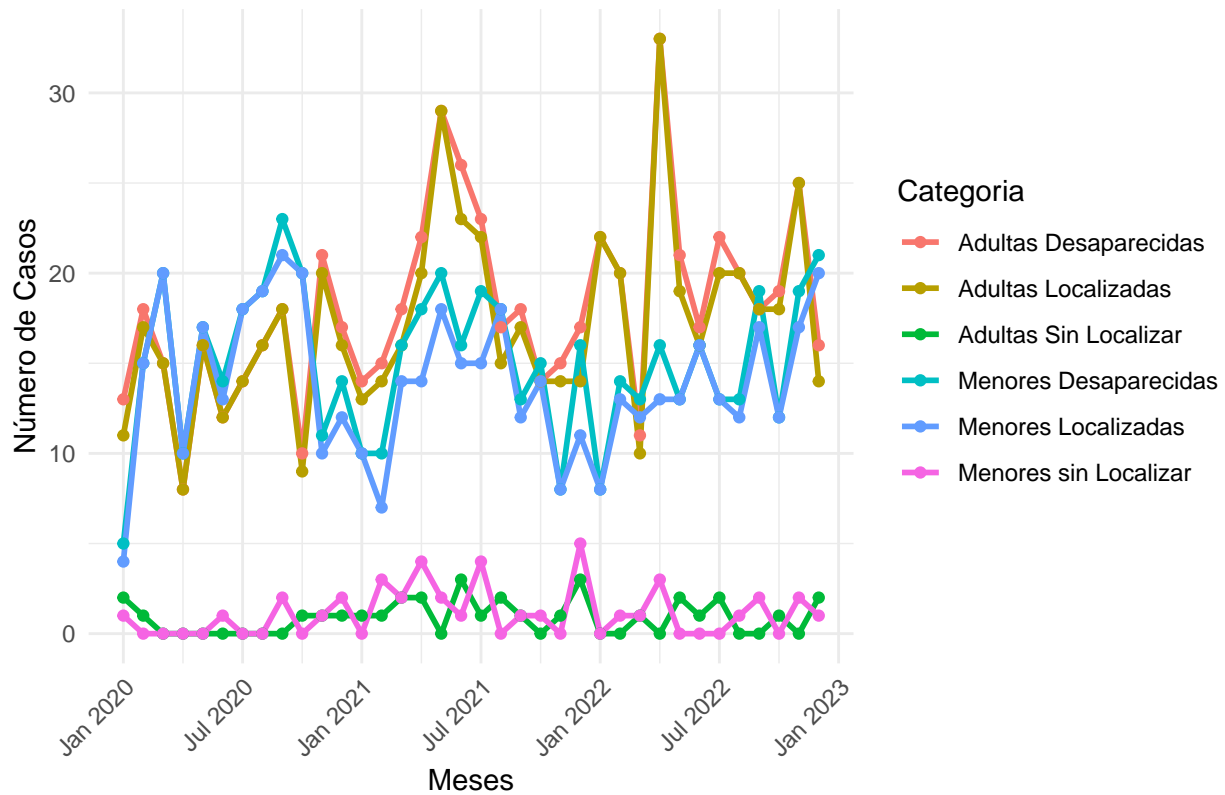
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.

## i Please use `linewidth` instead.

## 2: Tendencia de Casos por Mes y Categoría

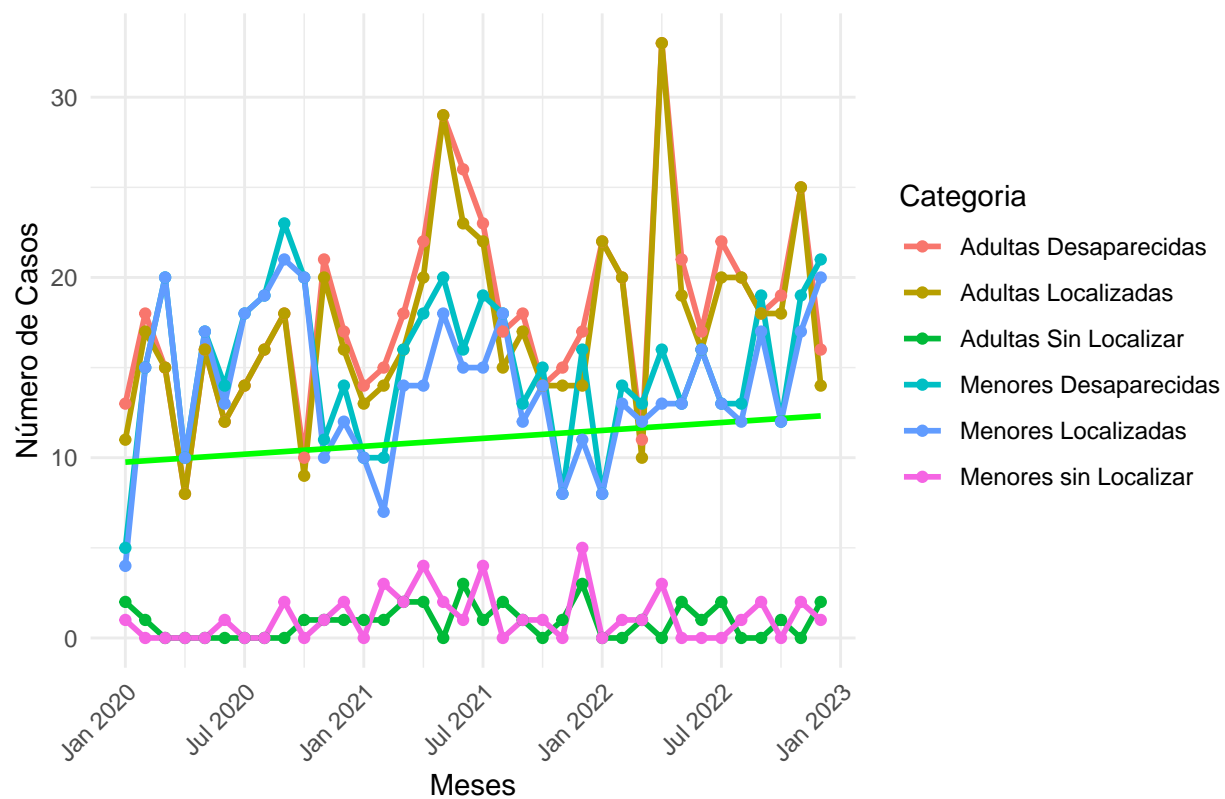


## 3: Tendencia de Casos por Mes y Categoría



## `geom\_smooth()` using formula = 'y ~ x'

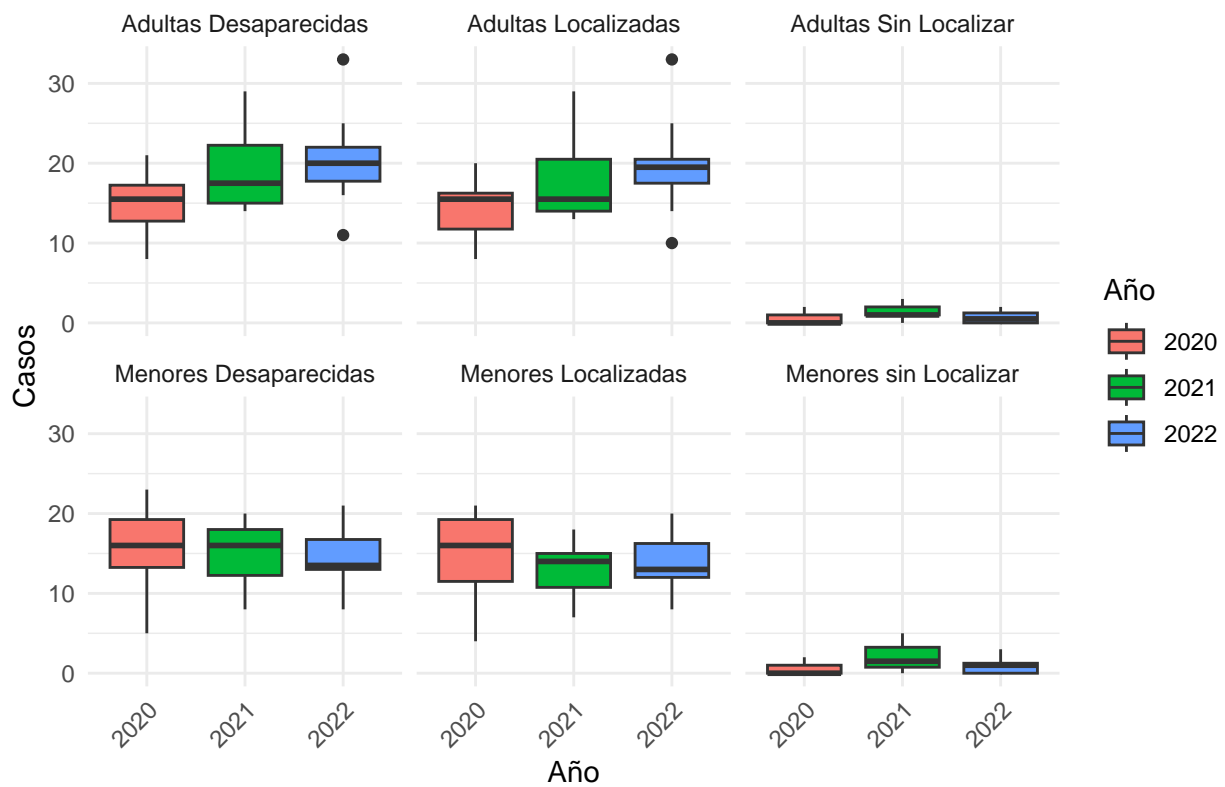
#### 4: Tendencia de Casos por Mes y Categoría



Muestra como varían los Casos a lo largo del tiempo para cada grupo.

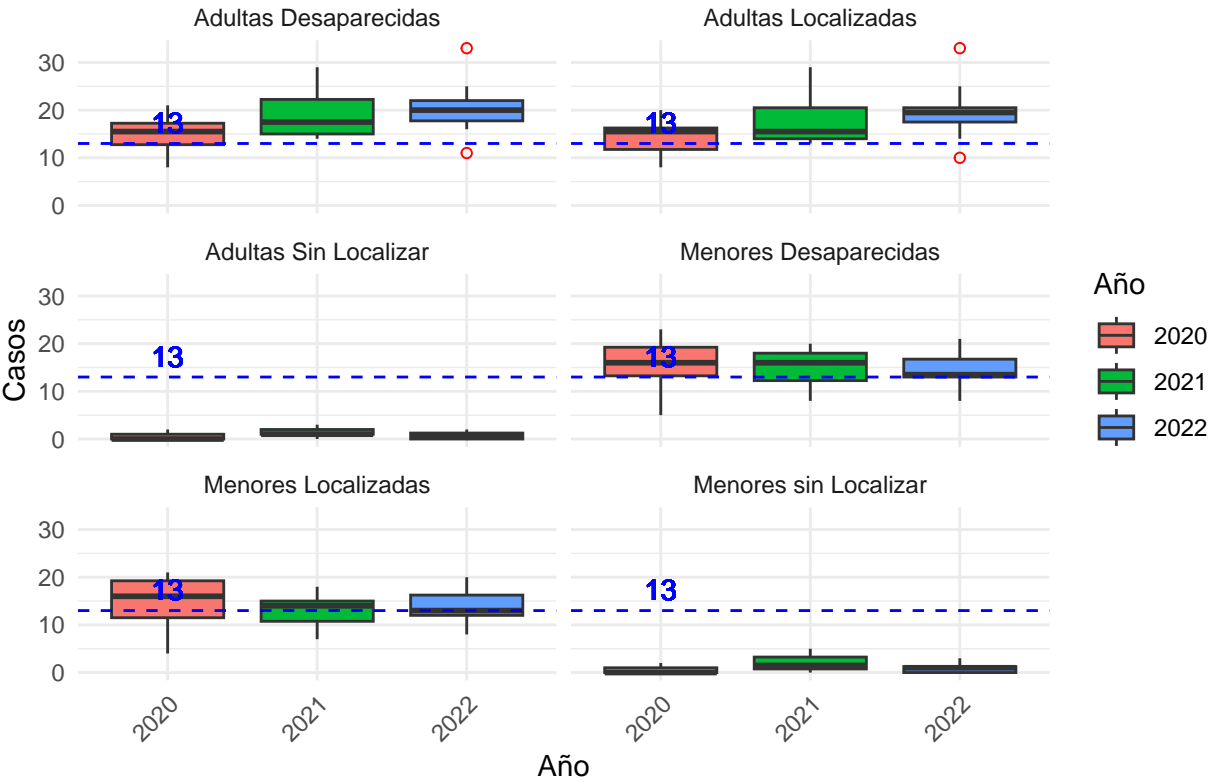
# Boxplot de la distribución de Casos por Año y Categoría

## 5: Distribución de Casos por Año y Categoría



```
## Warning: Use of `desp_long_xtotals$Casos` is discouraged.  
## i Use `Casos` instead.
```

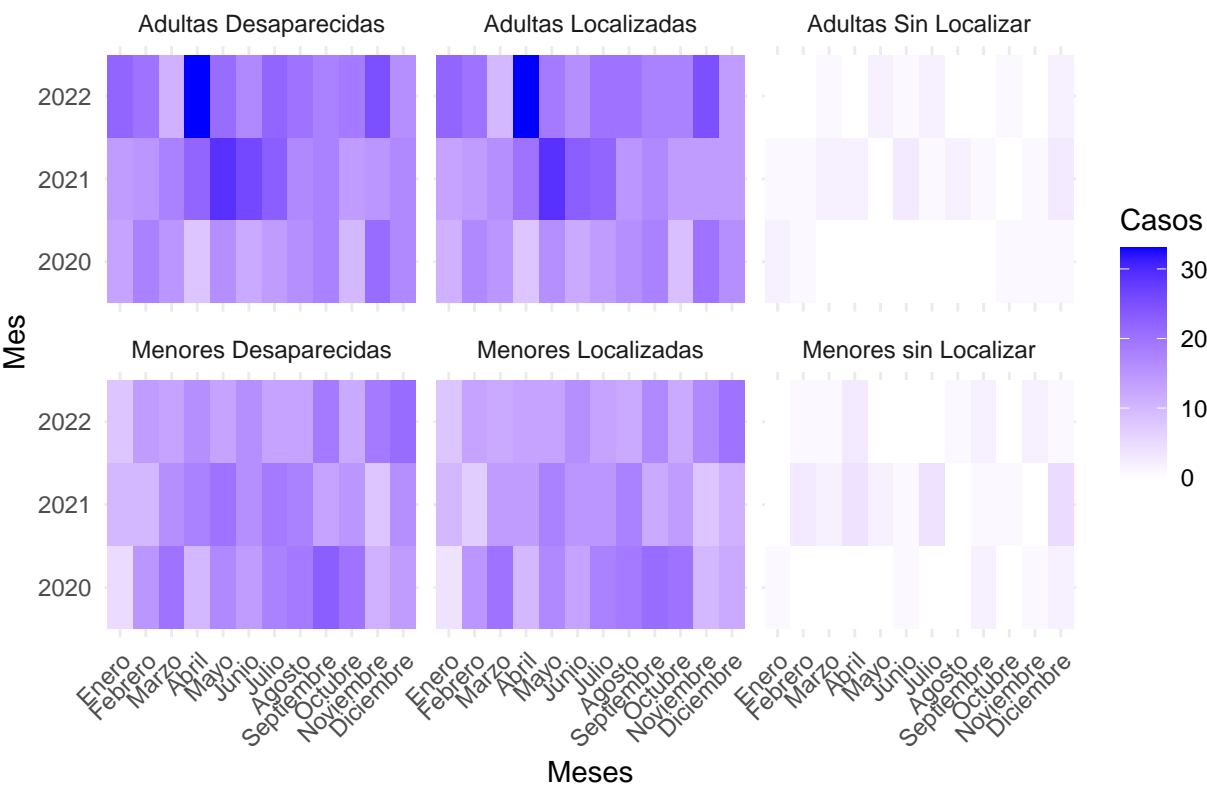
# 6: Distribución de Casos por Año y Categoría



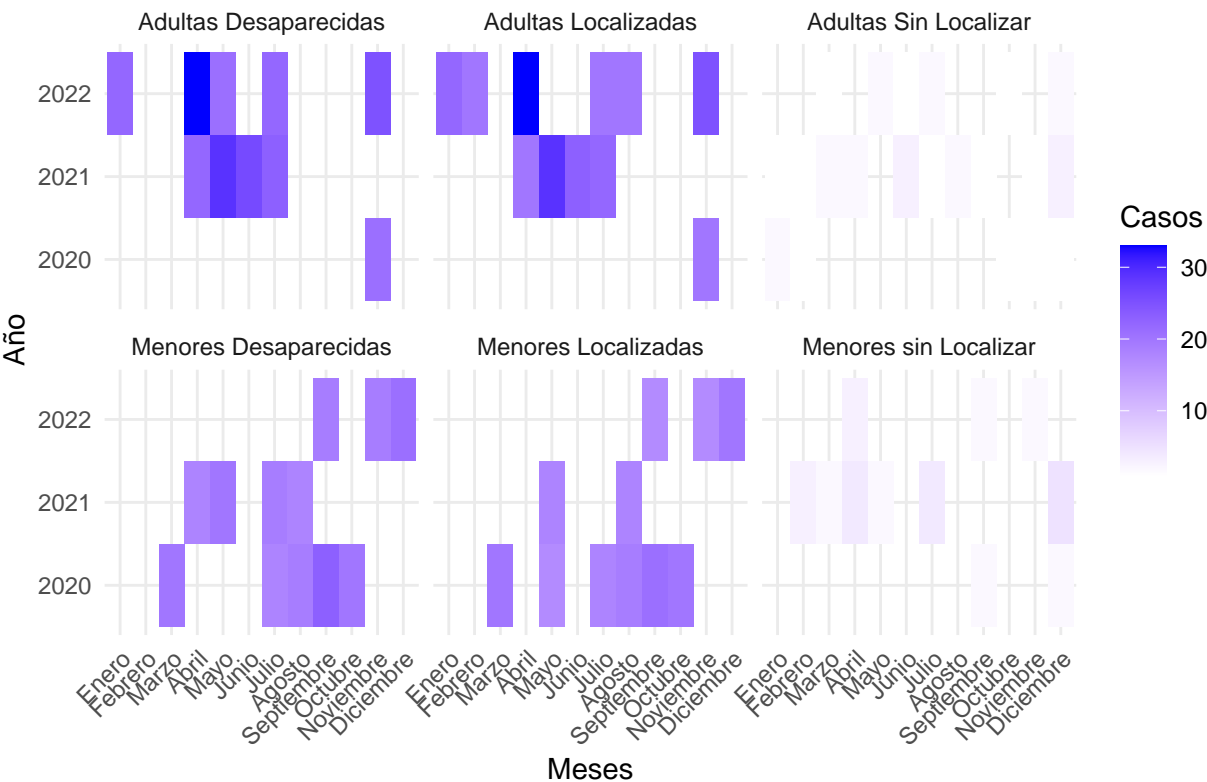
Muestra la distribución de Casos en cada mes para las diferentes categorías.

# Mapa de Calor de Casos por Mes y Categoría

7: Distribución de Casos por Mes y Año, Agrupados por Categoría



8: Mapa de Calor de los 10 Segmentos Principales por Mes y Categoría

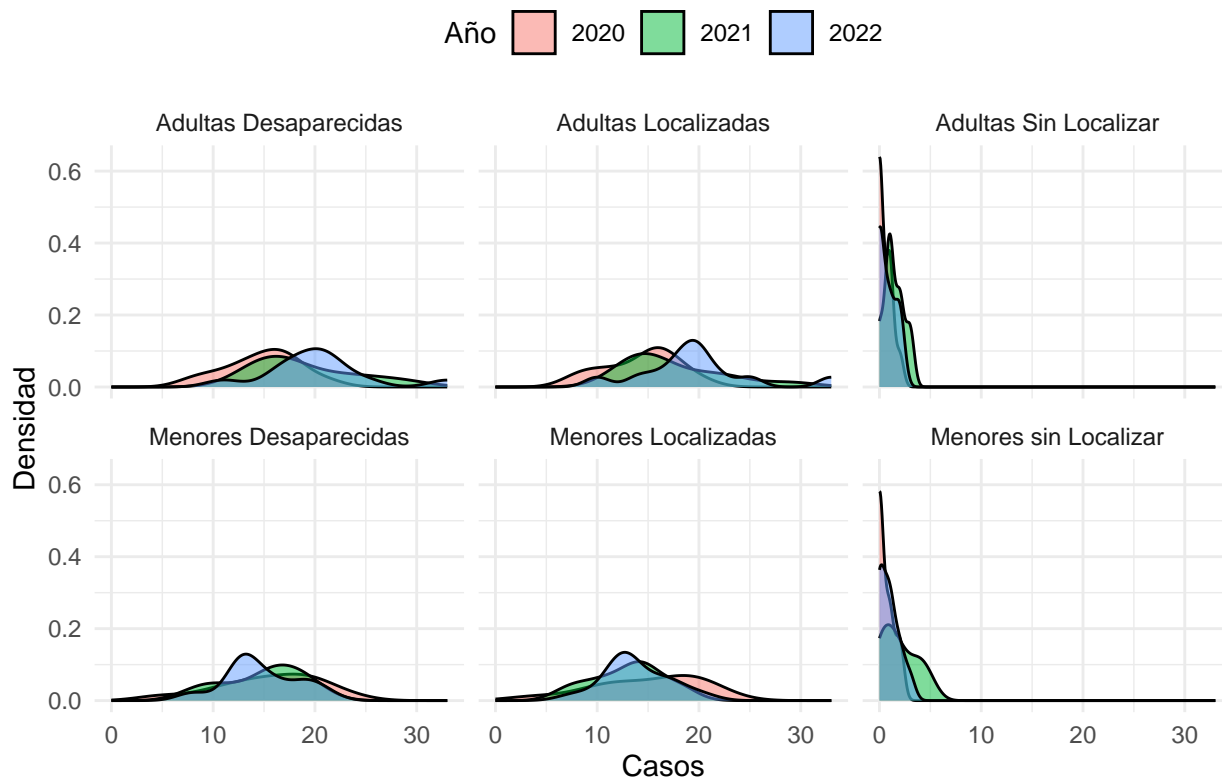


Muestra la diferencia entre Casos en diferentes meses para cada categoría.

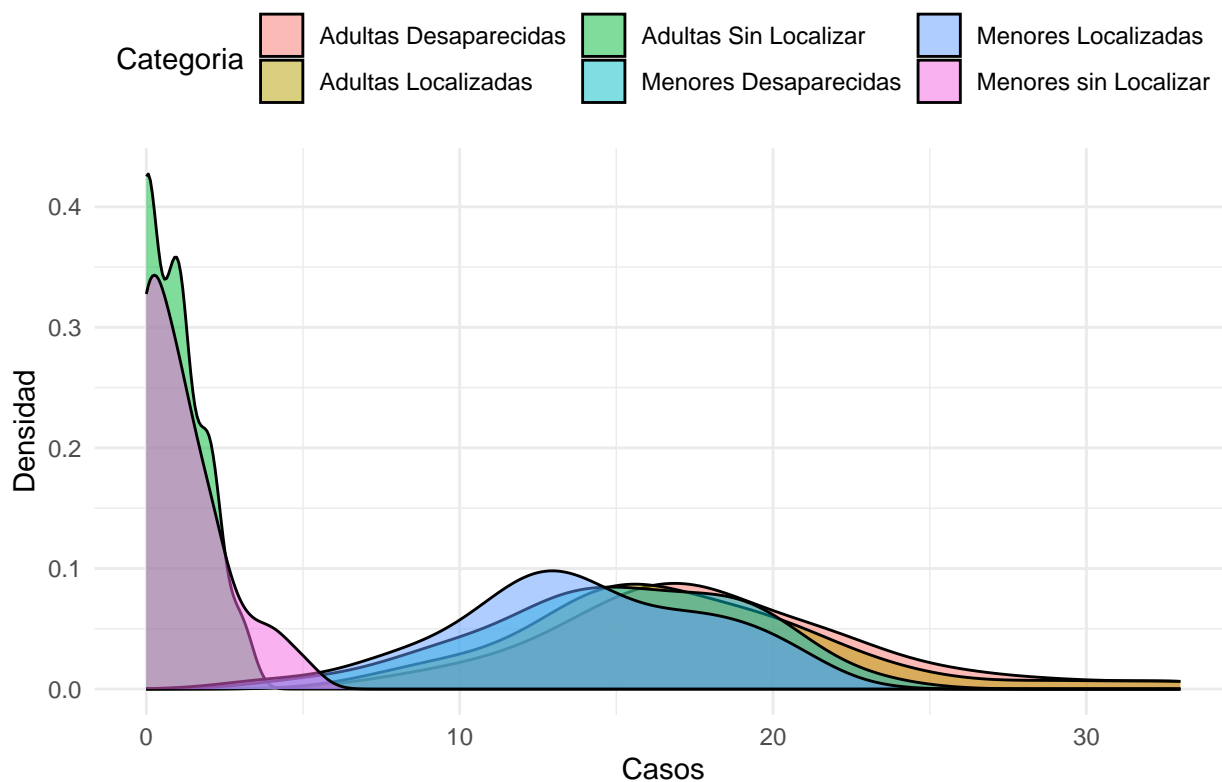


## Grafico de Densidad

### 9: Densidad de Casos por Categoría



### 10: Densidad de Casos por Categoría



Muestra la distribución de los Casos para cada categoría en cada mes usando curvas de densidad.

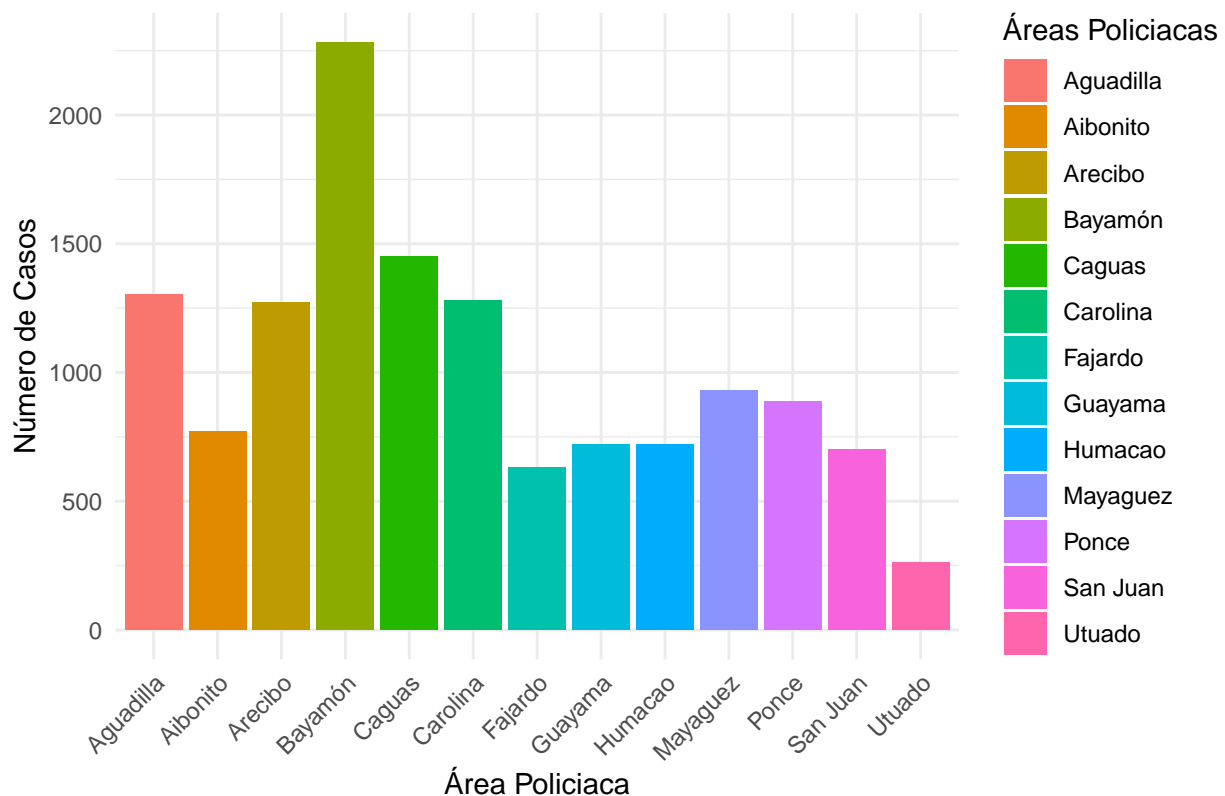
Muestra cómo cambian los Casos a lo largo de los meses para cada año, permitiendo comparar las tendencias a lo largo del tiempo.

## Incidentes de violencia doméstica por edad de la víctima (desde enero 2021 a abril 2023)

```
## # A tibble: 432 x 4
##   `Áreas Policiacas` Población Casos Fecha
##   <chr>              <dbl> <dbl> <yearmon>
## 1 Aguadilla          227732    44 Jan 2021
## 2 Aguadilla          227732    55 Feb 2021
## 3 Aguadilla          227732    72 Mar 2021
## 4 Aguadilla          227732    44 Apr 2021
## 5 Aguadilla          227732    75 May 2021
## 6 Aguadilla          227732    43 Jun 2021
## 7 Aguadilla          227732    33 Jul 2021
## 8 Aguadilla          227732    45 Aug 2021
## 9 Aguadilla          227732    32 Sep 2021
## 10 Aguadilla         227732    62 Oct 2021
## # ... with 422 more rows

## Warning: Removed 104 rows containing missing values (`position_stack()`).
```

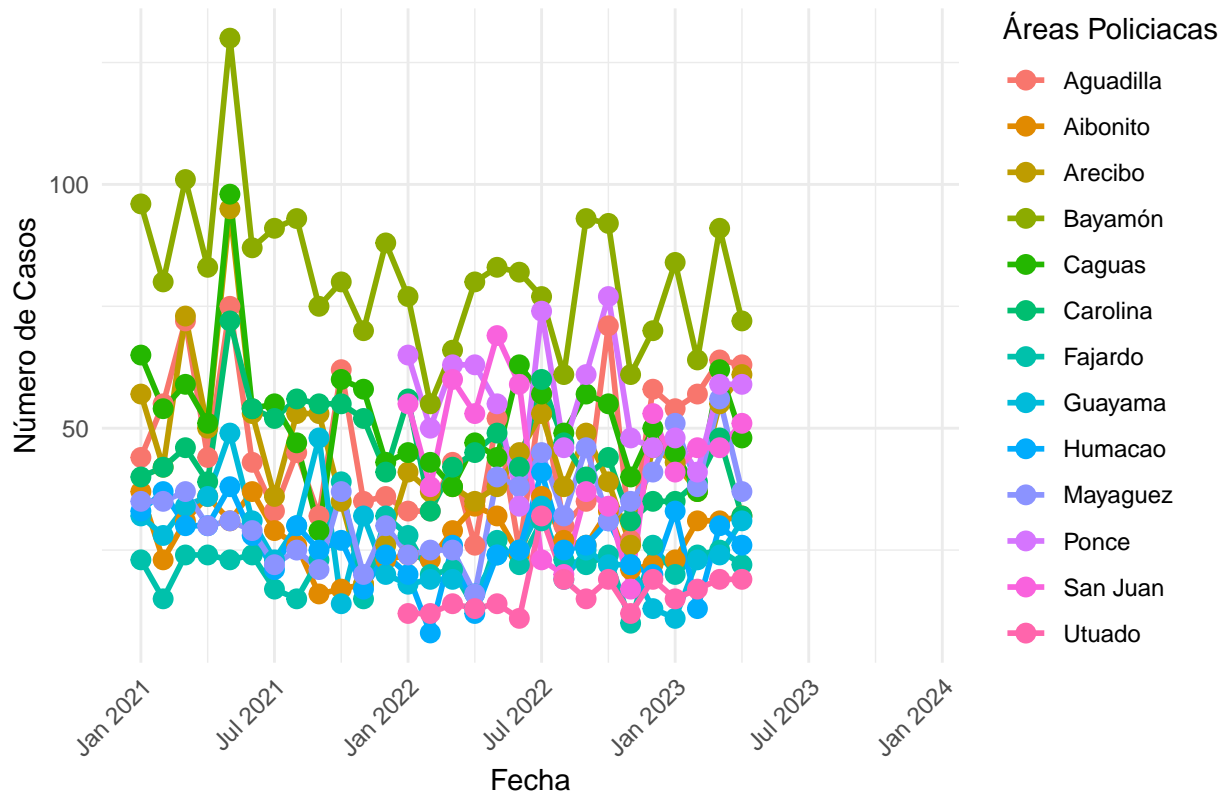
### 11: Distribución de Casos por Área Policiaca



```
## Warning: Removed 104 rows containing missing values (`geom_line()`).
```

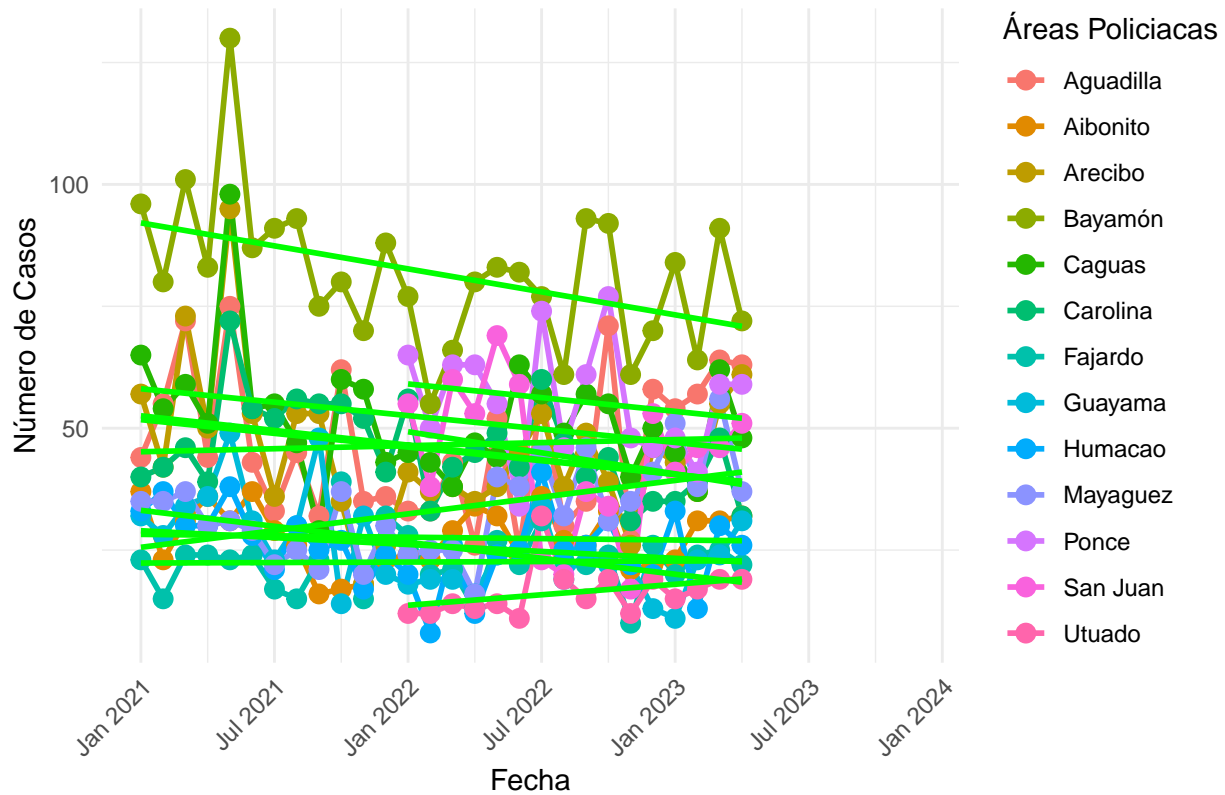
```
## Warning: Removed 104 rows containing missing values (`geom_point()`).
```

## 12: Tendencia de Casos a lo Largo del Tiempo por Área Policiaca



```
## `geom_smooth()` using formula = 'y ~ x'
## Warning: Removed 104 rows containing non-finite values (`stat_smooth()`).
## Warning: Removed 104 rows containing missing values (`geom_line()`).
## Warning: Removed 104 rows containing missing values (`geom_point()`).
```

### 13: Tendencia de Casos a lo Largo del Tiempo por Área Policiaca

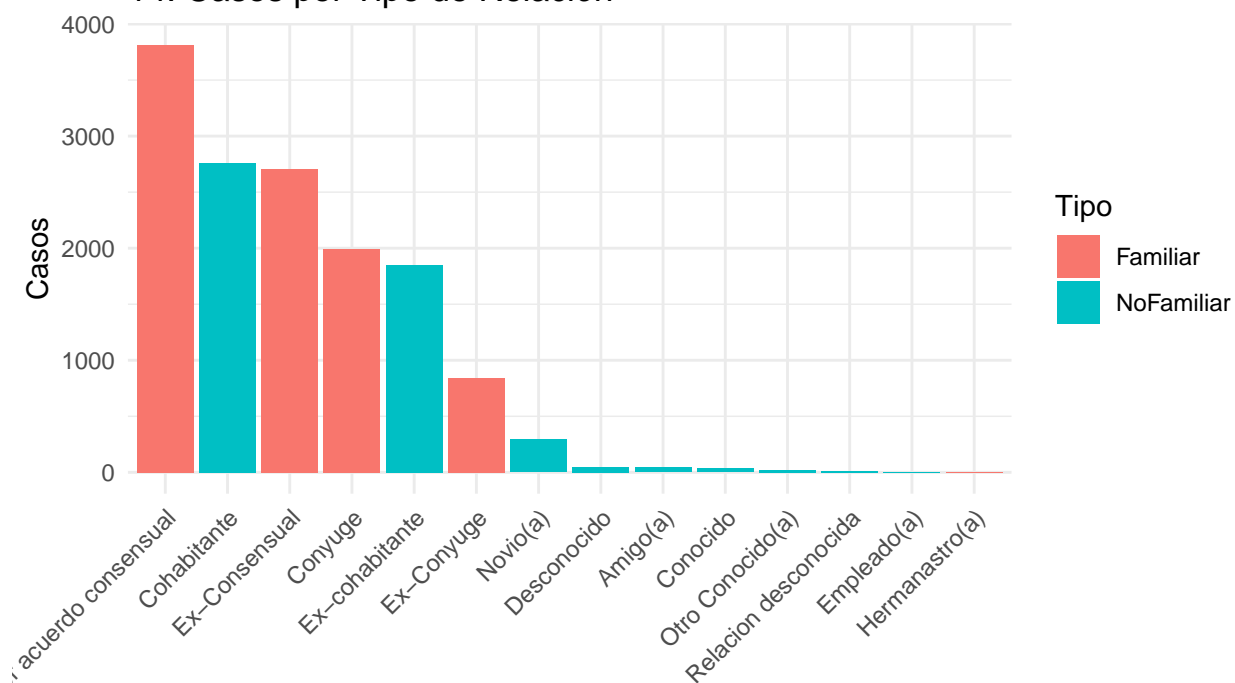


### Relación ofensor-víctima violencia doméstica (desde enero 2021 a febrero de 2023)

```
## # A tibble: 14 x 3
##   Relación                Tipo      Casos
##   <chr>                   <chr>    <dbl>
## 1 Conyuge por acuerdo consensual Familiar    3815
## 2 Ex-Consensual           Familiar    2705
## 3 Conyuge                 Familiar    1991
## 4 Ex-Conyuge              Familiar     842
## 5 Hermanastro(a)         Familiar         1
## 6 Cohabitante            NoFamiliar   2760
## 7 Ex-cohabitante         NoFamiliar   1852
## 8 Novio(a)               NoFamiliar    294
## 9 Desconocido            NoFamiliar     49
## 10 Amigo(a)              NoFamiliar     43
## 11 Conocido              NoFamiliar     34
## 12 Otro Conocido(a)      NoFamiliar     22
## 13 Relacion desconocida  NoFamiliar     11
## 14 Empleado(a)          NoFamiliar      1

## # A tibble: 2 x 5
##   Tipo      Total_Casos Promedio_Casos Casos_Mayor Casos_Menor
##   <chr>         <dbl>         <dbl>         <dbl>         <dbl>
## 1 Familiar      9354         1871.         3815           1
## 2 NoFamiliar    5066          563.         2760           1
```

14: Casos por Tipo de Relación



Tipo de Relación

15: Tendencia de Casos por Tipo y Relación

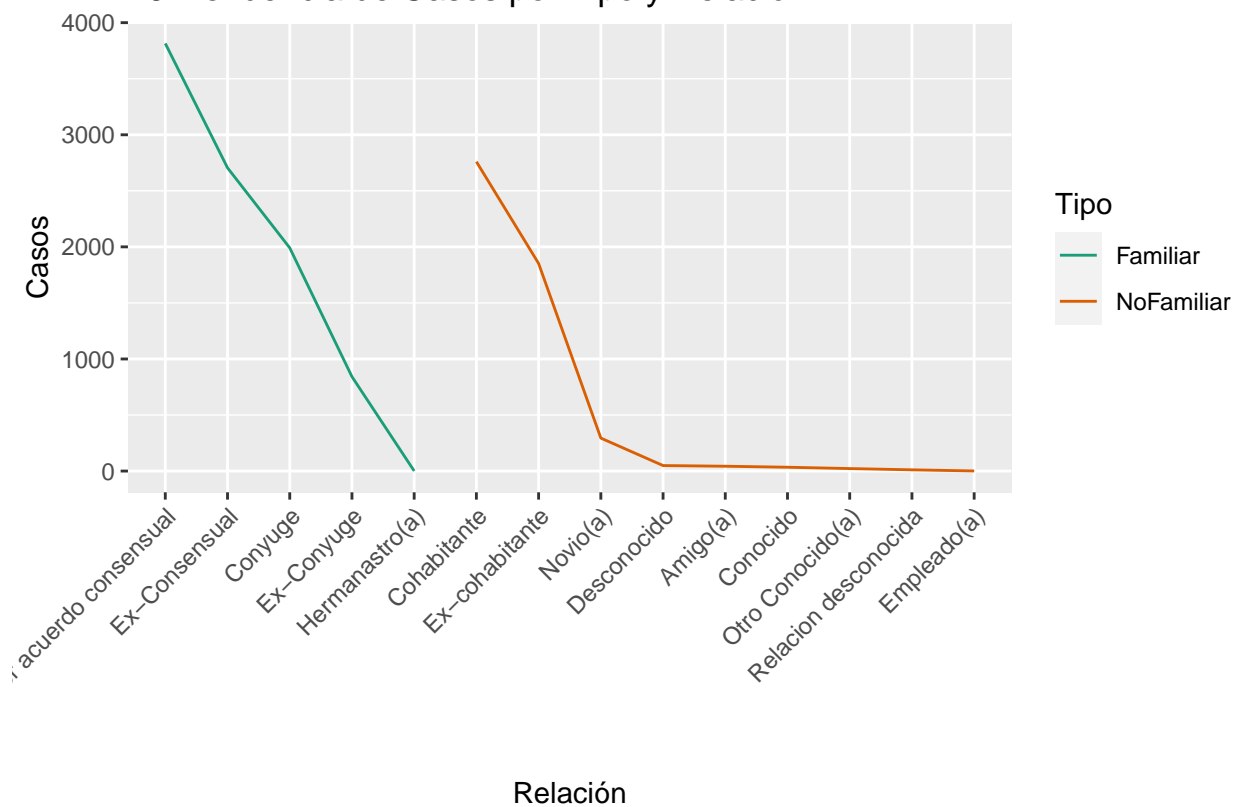
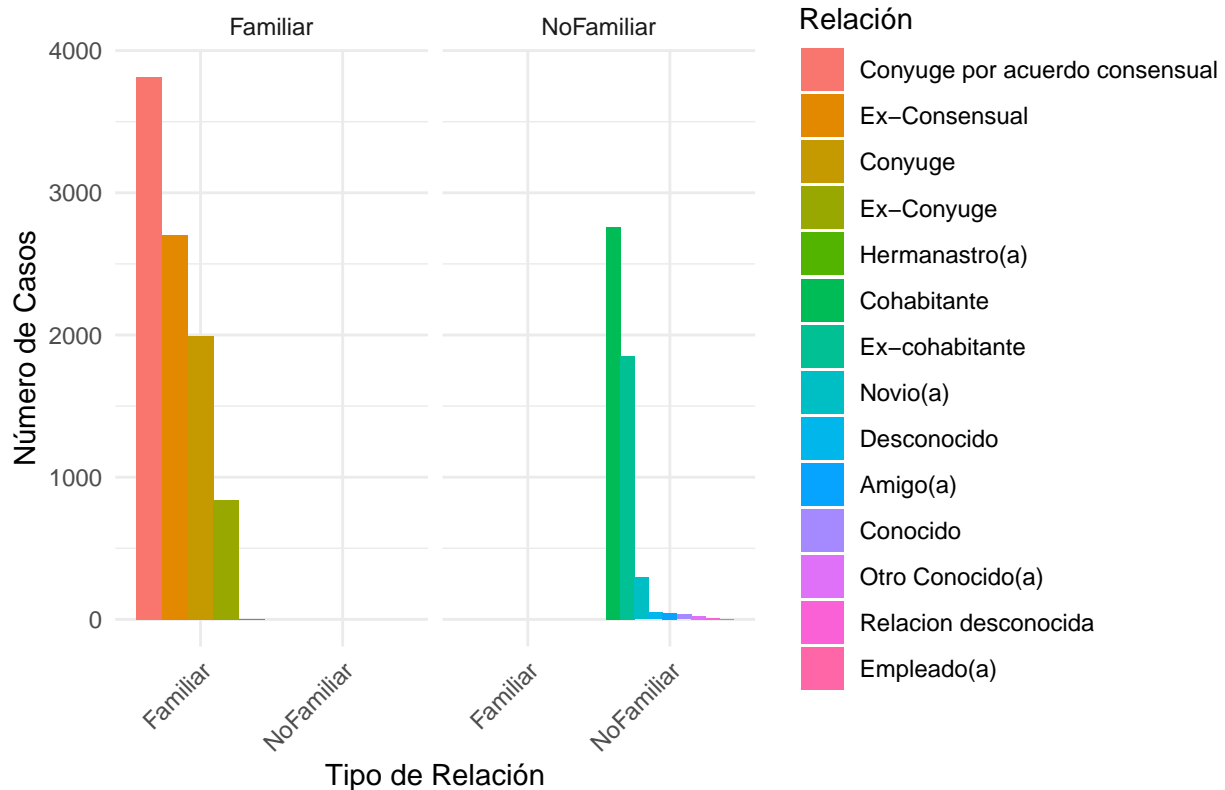


Table 2: Resumen estadístico por tipo de relación

Categoría	Media	Mediana	Desviación Estándar
Adultas Desaparecidas	18.0555556	17.5	5.1377841
Adultas Localizadas	17.1666667	16.0	5.1796304
Adultas Sin Localizar	0.8888889	1.0	0.9189366
Menores Desaparecidas	15.0555556	15.5	4.1261843
Menores Localizadas	13.9166667	13.5	4.0240349
Menores sin Localizar	1.1388889	1.0	1.3342259

## 16: Casos por Relación, Agrupados por Tipo



## T-Test entre Familiar y No Familiar

Las pruebas t se utilizan para comparar si las medias de dos grupos son significativamente diferentes.

```
##
## Welch Two Sample t-test
##
## data: familiar and no_familiar
## t = 1.7364, df = 6.098, p-value = 0.1324
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -528.025 3143.847
## sample estimates:
## mean of x mean of y
## 1870.8000 562.8889
```

## Wilcoxon-Mann-Whitney entre Familiar y No Familiar

Esta prueba se utiliza para verificar si las medianas de los grupos Familiar y NoFamiliar son significativamente diferentes

```
## Warning in wilcox.test.default(familiar, no_familiar): cannot compute exact
## p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: familiar and no_familiar
## W = 32.5, p-value = 0.2048
## alternative hypothesis: true location shift is not equal to 0
```

## ANOVA entre Familiar y No Familiar

Se usa para comparar las medias entre Familiar y No Familiar

```
## Analysis of Variance Table
##
## Response: Casos
##           Df    Sum Sq Mean Sq F value    Pr(>F)
## Tipo       1  5498458 5498458   3.8067 0.07479 .
## Residuals 12 17333190 1444432
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

## Chi-Squared Test entre Familiar y No Familiar

Esta prueba se utiliza para determinar si existe una asociación entre dos variables categóricas.

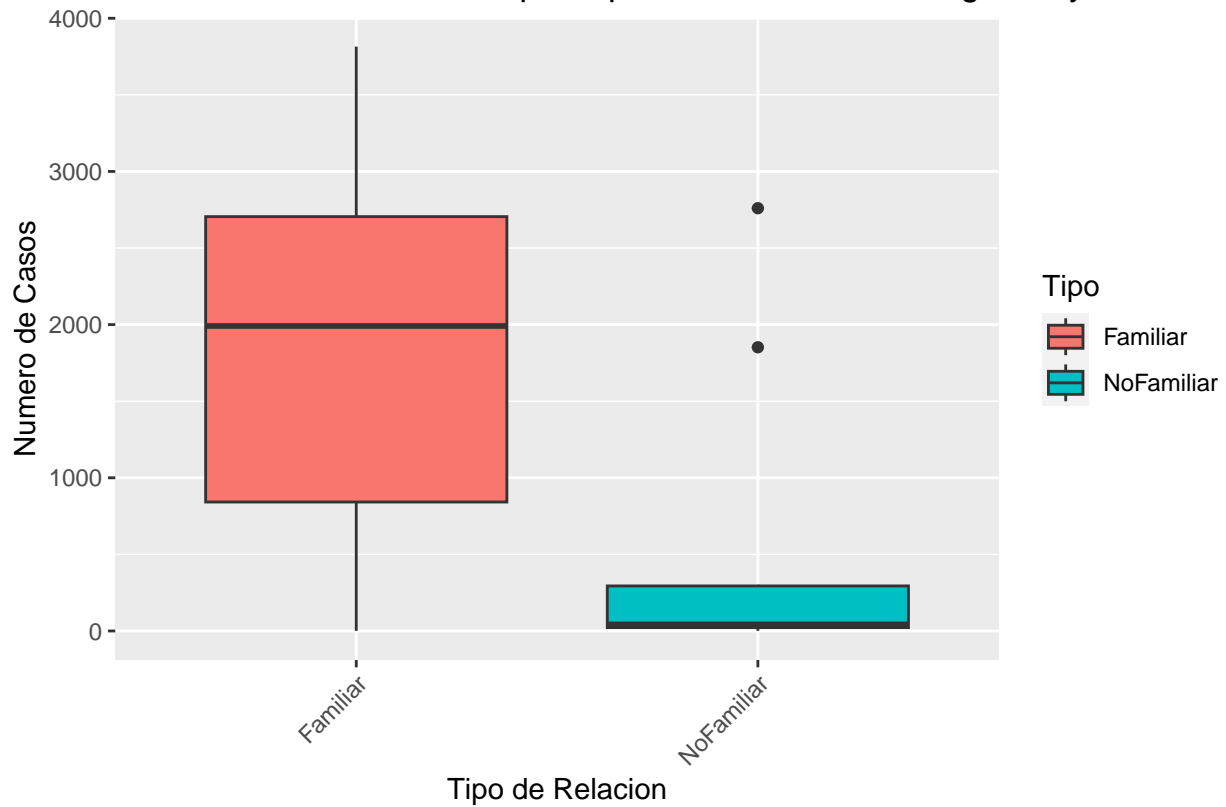
```
## Warning in chisq.test(tabla_contingencia): Chi-squared approximation may be
## incorrect

##
## Pearson's Chi-squared test
##
## data: tabla_contingencia
## X-squared = 14, df = 13, p-value = 0.3738

## Warning in kable_styling(.): Please specify format in kable.
## kableExtra can customize either HTML or LaTeX outputs. See
## https://haozhu233.github.io/kableExtra/ for details.
```

Prueba Estadística	Valor p
Wilcoxon-Mann-Whitney	0.2047762
Prueba t	0.1323758
ANOVA	0.0747888
Chi-cuadrado	0.3738440

17: Distribución de Casos por Tipo de Relacion entre Agresor y Victima



18: Numero de casos por tipo de relación ofensor-victima

