

Análisis exploratorio de la base de datos de la línea 144

Autor: Gonzalo Gabriel Ramirez

Fecha: 24/6/2023

Fuente: [Ministerio de las Mujeres, Géneros y Diversidad de la Nación - Dirección Técnica de Registros y Bases de Datos](#).

1. Introducción

En el año 2013 se creó en la República Argentina la línea 144. Dicha línea de comunicación telefónica nace con la intención de cumplir los objetivos establecidos por la Ley 26.845, artículo 9. Su tarea es la de brindar atención, contención y asesoramiento a mujeres y LGTBI+ en situación de violencia de género. Su equipo interdisciplinario está compuesto por profesionales de las áreas del derecho, la psicología, el trabajo social y otras áreas afines con capacitación en perspectiva de género, según informa [el sitio web oficial del Gobierno de la República Argentina](#).

Nuestra tarea en este análisis será manipular y presentar en gráficos estos datos para responder a distintas cuestiones acerca de la línea 144. Como por ejemplo *¿cuál es la distribución de las provincias de residencia de las personas que denuncian algún tipo de violencia?* o *¿cuál es la edad media de las personas que denuncian sufrir violencia de género en cualquiera de sus tipos?*. Éstas y otras cuestiones serán presentadas, analizadas y respondidas en base a la información que nuestro dataframe nos ofrece.

2. El dataframe

Nuestro dataframe es el resultado de la concatenación de 4 archivos de tipo '.csv' provisto por el sitio web [Datos.gov.ar](#). Cada archivo '.csv' corresponde al período 2020, 2021, 2022 y al primer trimestre del 2023 respectivamente. La sumatoria de estos 4 periodos nos dá como resultado **84.861 observaciones** y 18 variables:

- **fecha:** - ISO-8601 (date) - Fecha del ingreso de la consulta.
- **prov_residencia_persona_en_situacion_violencia:** - Texto (string) - Provincia en donde la persona en situación de violencia declara estar residiendo al momento de la consulta.
- **genero_persona_en_situacion_de_violencia:** - Texto (string) - Identidad de género autopercebida de la persona en situación de violencia.
- **edad_persona_en_situacion_de_violencia Texto** - Texto (string) - Edad declarada de la persona en situación de violencia al momento en que se ingresó la consulta.
- **pais_nacimiento_persona_en_situacion_de_violencia** - Texto (string) - País de nacimiento de la persona en situación de violencia.

- **tipo_de_violencia_fisica** - Texto (string) - Tipo de violencia, contemplado por la ley 26.485, que motiva la consulta.
- **tipo_de_violencia_psicologica** - Texto (string) - Tipo de violencia, contemplado por la ley 26.485, que motiva la consulta.
- **tipo_de_violencia_sexual** - Texto (string) - Tipo de violencia, contemplado por la ley 26.485, que motiva la consulta.
- **tipo_de_violencia_economica_y_patrimonial** - Texto (string) - Tipo de violencia, contemplado por la ley 26.485, que motiva la consulta.
- **tipo_de_violencia_simbolica Texto** - Texto (string) - Tipo de violencia, contemplado por la ley 26.485, que motiva la consulta.
- **tipo_de_violencia_domestica** - Texto (string) - Modalidad de violencia, contempladas por la ley 26.485, que motivan la consulta bajo la cual se desarrollaron.
- **modalidad_de_violencia_institucional** - Texto (string) - Modalidad de violencia, contempladas por la ley 26.485, que motivan la consulta bajo la cual se desarrollaron.
- **modalidad_de_violencia_laboral** - Texto (string) - Modalidad de violencia, contempladas por la ley 26.485, que motivan la consulta bajo la cual se desarrollaron.
- **modalidad_violencia_contra_libertad_reproductiva** Texto (string) - Modalidad de violencia, contempladas por la ley 26.485, que motivan la consulta bajo la cual se desarrollaron.
- **modalidad_de_violencia_obstetrica** - Texto (string) - Modalidad de violencia, contempladas por la ley 26.485, que motivan la consulta bajo la cual se desarrollaron.
- **modalidad_de_violencia_mediatica** - Texto (string) - Modalidad de violencia, contempladas por la ley 26.485, que motivan la consulta bajo la cual se desarrollaron.
- **modalidad_de_violencia_otras** - Texto (string) - Categoría "Otras modalidades" para dejar registro de la presencia de otras modalidades distintas a las mencionadas (por ejemplo, violencia política, acoso callejero).
- **vinculo_con_la_persona_agresora** - Texto (string) - Vínculo que tiene o tenía la persona en situación de violencia con quien ejerce la agresión en el momento en que sucedieron las situaciones de violencia declaradas.
- **genero_de_la_persona_agresora** - Texto (string) - Identidad de género de la persona agresora.


Mediante la manipulación de los datos con la librería *pandas* se logró reunir y almacenar en una sola variable llamada *df* la totalidad de los registros de los cuatros periodos en cuestión, de la cual se desprenderá las distintas operaciones necesarias para responder cada una de las preguntas que abordaremos en este análisis.

Cabe destacar que nuestro dataframe presenta valores nulos en más de una variable. Todos los valores nulos fueron oportunamente sustituidos por la leyenda "Missing" que nos indica que ese valor no fue cargado desde su origen y se tendrá en cuenta para su descarte de acuerdo a la necesidad del momento o de la pregunta a responder.

Importamos librerías

Cargamos nuestros archivos .csv

	fecha	prov_residencia_persona_en_situacion_violencia	genero_persona_en_situacion_de_violencia
0	2020-01-01	Tucumán	Mujer
1	2020-01-01	Buenos Aires	Mujer
2	2020-01-01	Buenos Aires	Mujer
3	2020-01-01	Buenos Aires	Mujer
4	2020-01-01	Buenos Aires	Mujer



3. La edad y su distribución

En esta sección exploraremos dos preguntas relacionadas con la edad de los llamantes que solicitan asistencia a la línea 144:

- ¿Cómo se distribuye la edad de las personas que llaman?
- ¿Cuál es la edad promedio de los solicitantes de asistencia en la línea 144?

Para responder a estas preguntas, analizaremos nuestro dataframe *df_edades* y nos apoyaremos en un gráfico de distribución. Al observar la gráfica, podemos notar un patrón interesante: el número de llamadas aumenta significativamente a partir de los 20 años, alcanza su máximo entre los 30 y 40 años y comienza a disminuir a partir de los 45 años aproximadamente.

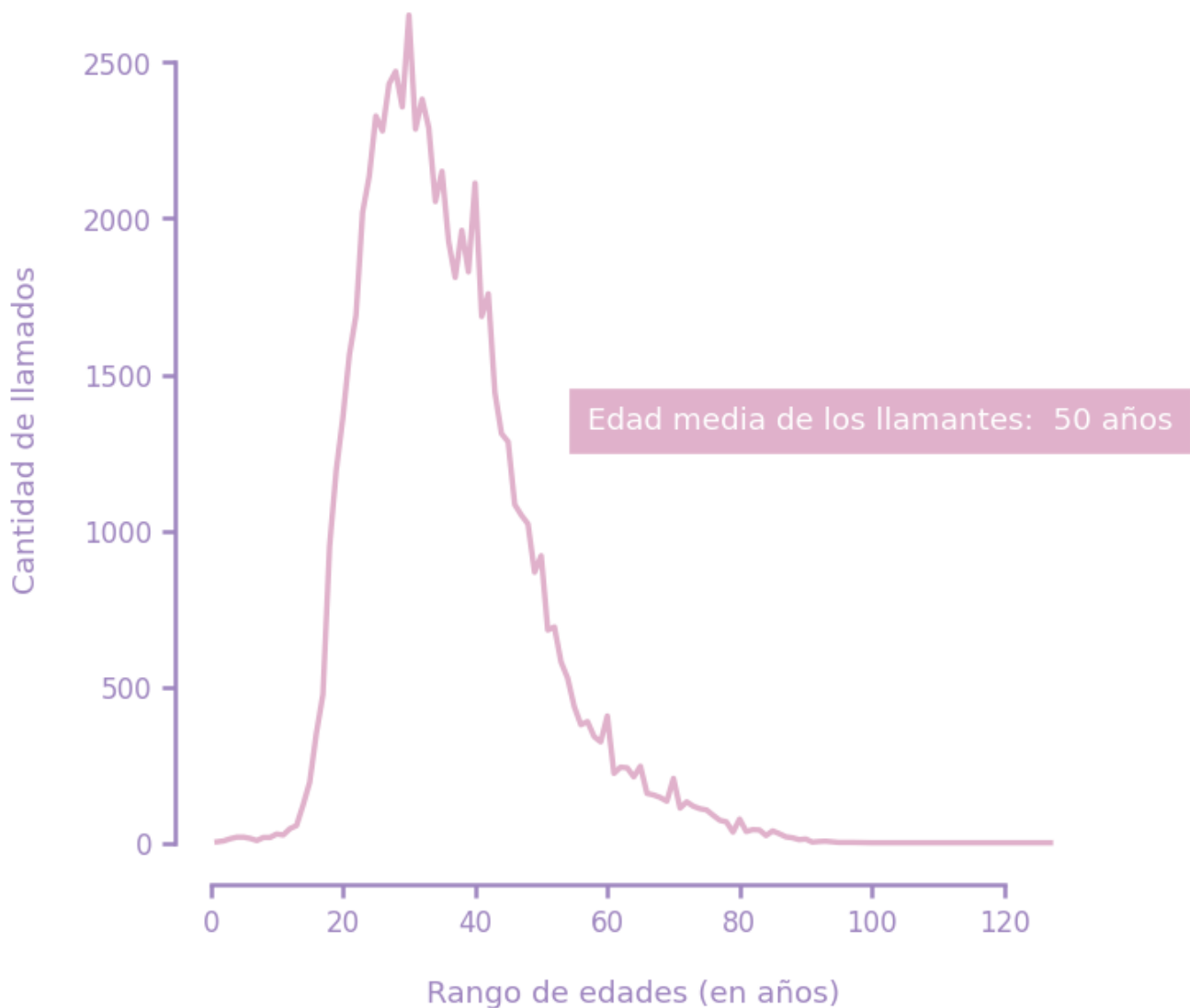
La edad media de los llamantes que denuncian en la línea 144 es de 36 años, lo que indica que el grupo de solicitantes de asistencia tiende a tener una edad cercana a esta cifra.

Estos hallazgos nos brindan una comprensión más profunda sobre cómo la edad está relacionada con las llamadas a la línea 144 y nos permite identificar el perfil predominante de quienes buscan ayuda en situaciones de violencia de género.

	Edad	Cantidad de llamadas
93	1.0	4
89	2.0	7
85	3.0	14
79	4.0	19
80	5.0	19
...
95	95.0	2
96	97.0	2
97	100.0	1
99	103.0	1
98	127.0	1

99 rows × 2 columns

50.0



4. El género

En esta sección, nos enfocaremos en analizar el género de las personas que solicitan asistencia a través de la línea 144. Cada individuo que realiza una llamada debe responder una serie de preguntas formuladas por el equipo interdisciplinario con el objetivo de llevar a cabo estadísticas y proporcionar la asistencia adecuada según la situación particular de cada individuo. Una de estas preguntas es sobre el género de la persona en situación de violencia.

Explorar la distribución de géneros declarados en las llamadas resulta relevante desde una perspectiva estadística para responder a las siguientes preguntas clave:

- ¿Cuál es la distribución de géneros declarados en las llamadas?
- ¿Predominan las mujeres como víctimas de algún tipo de violencia de género?

	Género	Conteo
0	Mujer	81204
2	Varon	326
3	Transgenero	139
4	Mujer Trans	128
5	Otro	61
6	Varon trans	24
7	Travesti	11
8	Intersexual	4
9	No Binarie	3
10	Queer	1

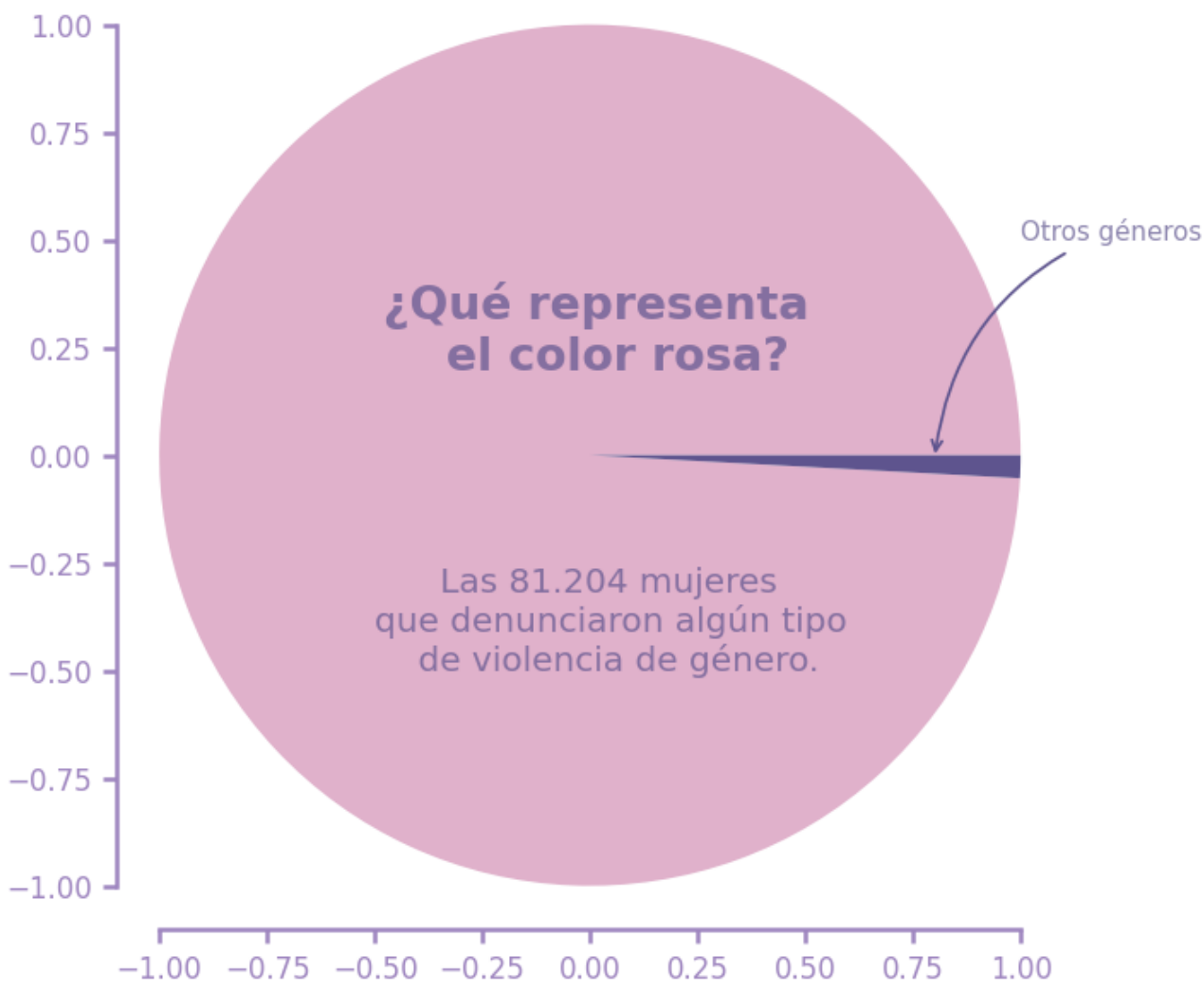
En este primer gráfico de torta, también conocido como piechart, examinaremos la distribución de géneros presentes en nuestro dataframe original. Encontramos 10 variables categóricas para el género, que son:

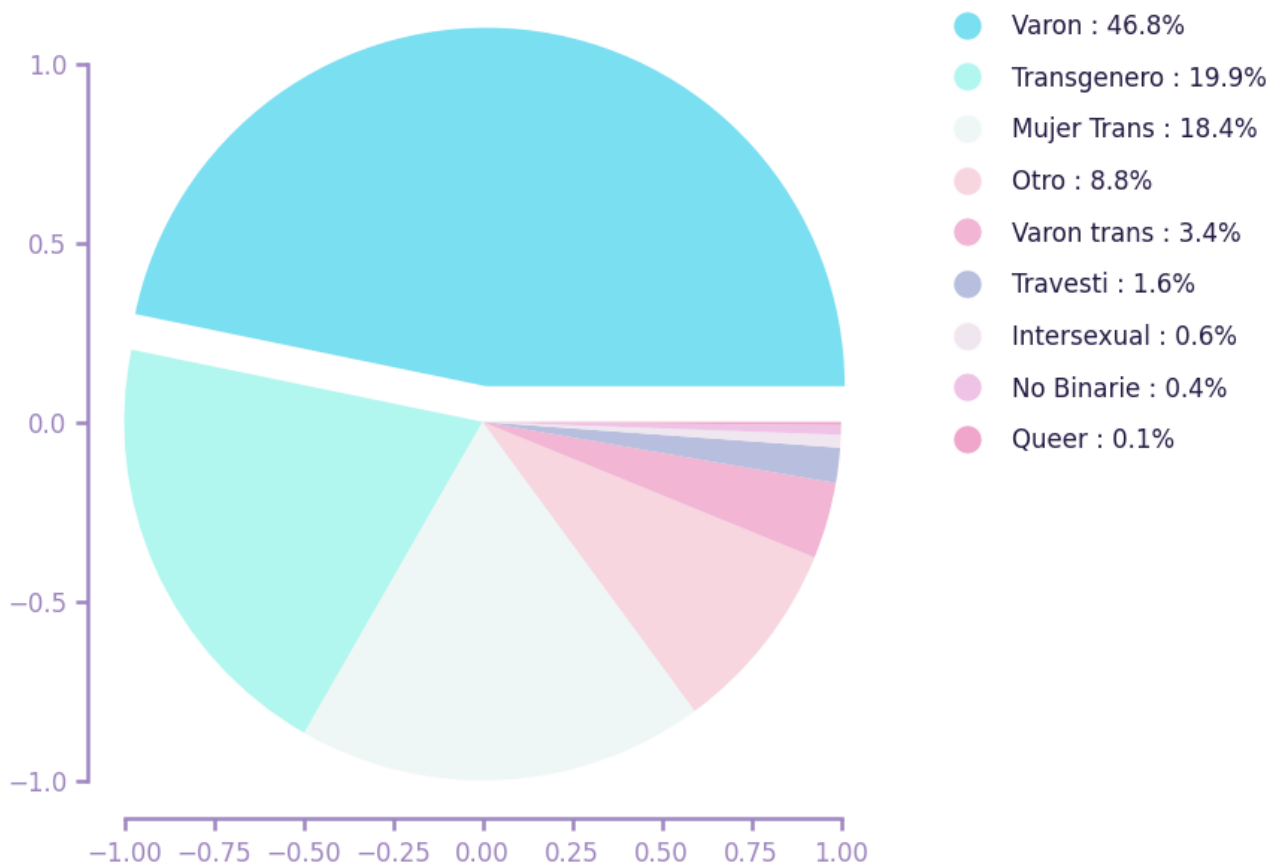
- Mujer
- Varón
- Mujer trans
- Transgénero
- Otro
- Varón trans
- Travesti
- Intersexual
- No binarie
- Queer

Al observar los datos de nuestro dataframe *generos*, se destaca una amplia diferencia numérica entre la categoría Mujer y las demás variables. Con el objetivo de representar claramente esta diferencia, hemos optado por agrupar todas las demás variables en una sola denominada *Otros*. Como resultado, operamos con dos variables: *Mujer* y *Otros*.

El piechart que se presenta a continuación nos permite dimensionar la relación entre la categoría Mujer y las demás. Aquí, a través de números y estadísticas descriptivas, podemos comprobar que de la totalidad de los llamados, 81.204 corresponden a denuncias de algún tipo de violencia de género por parte de mujeres. Mientras que en el resto de los llamados (697 llamadas) encontramos individuos que declaran un género distinto al de mujer.

Este análisis resalta la importancia de reconocer las diferencias en las llamadas y nos brinda información valiosa para comprender el alcance de la violencia de género en nuestra sociedad.





5. Distribución de llamados en el tiempo

En esta última sección, veremos la distribución de los llamados a lo largo de los períodos comprendidos entre el año 2020 y el primer trimestre del 2023.

Observando nuestra gráfica podemos asegurar que a través del tiempo la cantidad de denuncias que la línea 144 recibió fue disminuyendo. Es tal así, que en el período Abril 2020 fue el máximo histórico de nuestro dataframe. Mientras que en Diciembre del año 2022 se logra alcanzar el mínimo histórico de llamadas a la línea.

	Periodo	Llamados
0	2020-01-01	2433
1	2020-02-01	1968
2	2020-03-01	2495
3	2020-04-01	2937
4	2020-05-01	2821
5	2020-06-01	2395
6	2020-07-01	2596
7	2020-08-01	2605
8	2020-09-01	2243
9	2020-10-01	2251
10	2020-11-01	2310
11	2020-12-01	2197
12	2021-01-01	2088
13	2021-02-01	2141
14	2021-03-01	2278
15	2021-04-01	1998
16	2021-05-01	1995
17	2021-06-01	1844
18	2021-07-01	2070
19	2021-08-01	2186
20	2021-09-01	1926
21	2021-10-01	2103
22	2021-11-01	2120
23	2021-12-01	2187
24	2022-01-01	1995
25	2022-02-01	1941
26	2022-03-01	2391
27	2022-04-01	2101
28	2022-05-01	2005
29	2022-06-01	1803
30	2022-07-01	1846
31	2022-08-01	2074

	Periodo	Llamados
32	2022-09-01	2100
33	2022-10-01	2070
34	2022-11-01	2023
35	2022-12-01	1889
36	2023-01-01	1900
37	2023-02-01	1580
38	2023-03-01	1755

