Informe Final: Análisis de Datos sobre Violencia de Género en Argentina

Institución: CoderHouse

Elaborado por: Marcos Praga

Fecha: junio de 2025

Contenido

Objetivo y alcance del proyecto	3
Objetivos del análisis:	4
Usuario final y nivel de aplicación del análisis	4
Herramientas tecnológicas utilizadas	4
Proceso de limpieza del dataset	5
Descripción de las tablas	6
Tabla Registro_Violencia	6
Tabla Genero Agresor:	7
Tabla Vinculo	7
Tabla País Nacimiento	7
Tabla Genero	8
Tabla Provincia	8
Modelo de entidad relación	9
Proceso de transformación de datos	10
Detalle de Solapas	14
Caratula	14
Glosario	15
Análisis General	16
Análisis Regional	18
Análisis Histórico	20
Conclusión	22

Objetivo y alcance del proyecto

El objetivo de este proyecto es desarrollar un tablero interactivo que consolide, visualice y analice los datos normalizados sobre violencia de género. A través de este tablero, buscamos ofrecer una herramienta clara, dinámica y accesible que permita a los usuarios finales interpretar tendencias, identificar patrones críticos y tomar decisiones informadas para la implementación de medidas de prevención, asistencia o políticas públicas específicas.

El proyecto comprende la creación de un tablero de visualización que represente de manera gráfica la información contenida en las tablas normalizadas de la base de datos. Este tablero incluirá indicadores clave, gráficos comparativos, filtros dinámicos y análisis segmentados por variables relevantes (como tipo de violencia, edad de las víctimas, localización geográfica, entre otros).

La capacidad de transformar datos crudos en información estratégica es fundamental en contextos donde se requiere intervención inmediata y eficiente. El norte de este estudio es que los datos obtenidos permitan a organismos públicos y ONGs visualizar tendencias, anticipar escenarios críticos y focalizar recursos.

La inteligencia de negocios aplicada a problemáticas sociales fortalece la transparencia, la planificación y la evaluación de impacto. En este sentido, este informe no solo es técnico, sino también político: contribuye a que las decisiones se basen en evidencia concreta y no en suposiciones.

Objetivos del análisis:

- Analizar la distribución geográfica de los casos de violencia de género en Argentina durante 2023.
- Determinar qué tipos de violencia son más frecuentes entre las personas afectadas.
- Analizar la relación entre la edad de la víctima y el tipo de vínculo con el agresor.
- Explorar si existe una diferencia en los tipos de violencia según el género del agresor.

Usuario final y nivel de aplicación del análisis

Usuario final: Entidades gubernamentales, organismos de justicia, ONGs, equipos de políticas públicas y analistas sociales que trabajen en la temática de violencia de género.

Nivel de aplicación táctico: está diseñado para facilitar la identificación de tendencias y focos de mayor incidencia en la violencia de género. Al permitir el análisis detallado de datos segmentados por variables clave como tipo de violencia, edad de las víctimas y ubicación geográfica, ofrece información precisa para la planificación de acciones concretas a corto y mediano plazo.

Las entidades gubernamentales, ONGs y equipos de políticas públicas pueden utilizar los resultados del análisis para ajustar estrategias de prevención, focalizar recursos y mejorar la eficiencia de intervenciones en áreas críticas. Al transformar datos en información visual y accesible, el dashboard no solo proporciona insights estratégicos, sino que también fortalece la toma de decisiones basada en evidencia en lugar de suposiciones.

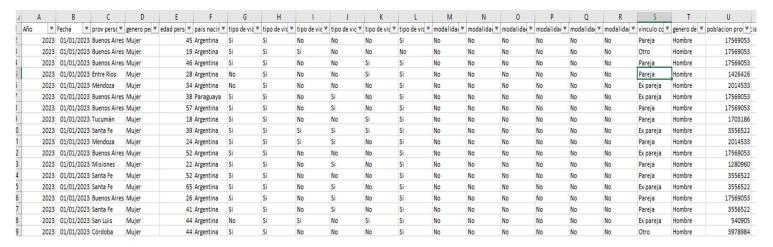
Herramientas tecnológicas utilizadas

Para la elaboración del trabajo final se utilizaron los siguientes programas:

- Excel para la lectura y limpieza de los datasets.
- Draw.io para la creación del diagrama entidad-relación
- Power BI Desktop para la creación del tablero de control.

Proceso de limpieza del dataset

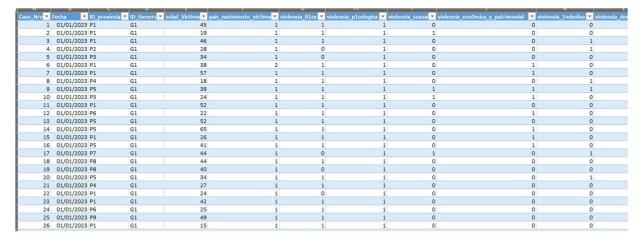
Al seleccionar el dataset de kaggle.com se encontraba de la siguiente manera:



Para normalizarlo procedimos a eliminar las filas vacías y convertir los sí y los no alojados en las celdas de los distintos tipos de violencia en 0 para los no y en 1 para los si con la fórmula de reemplazar desde Excel para poder sumar los casos que sí y dejar de lado los no. Además de esto se eliminó la columna de Año ya que la información estaba repetida en la columna fecha.

Agregamos también la columna de Caso_nro y le asignamos la clave primaria para la tabla registro_violencia. De esta manera lo hicimos incrementar para no tener números repetidos y poder identificar cada uno de los casos de manera individual.

Además de esto se eliminaron los espacios en los títulos de las columnas y por ultimo procedí a ordenar la información en distintas tablas y reemplazar los datos en la principal para poder convertir las llaves primarias y foráneas de las tablas. Quedando de la siguiente manera al insertar como tabla:



Descripción de las tablas

El conjunto de datos utilizado proviene de registros públicos anonimizados, estructurados en múltiples tablas normalizadas. La tabla principal es 'Registro_Violencia', donde se concentran los casos individuales y sus atributos: fecha, edad, género de víctima y agresor, tipos de violencia, y provincia de ocurrencia. El modelo relacional se completa con las tablas de dimensiones: 'Provincia', 'Género', 'País Nacimiento', 'Vínculo' y 'Género del agresor'.

Tabla Registro Violencia

Es la tabla principal que alberga todos los registros de los casos analizados con sus respectivos detalles:

Atributos:

Tabla Registro_violencia		
Campo	Tipo de campo	Tipo de clave
Caso_Nro	Int	PK
Fecha	date	-
ID provincia	varchar (50)	FK
ID_Genero	varchar (50)	FK
edad victima	Int	-
pais nacimiento victima	Int	FK
violencia fisica	Int	-
violencia_psicologica	Int	-
violencia_sexual	Int	-
violencia_economica_y_patrimonial	Int	-
violencia_simbolica	Int	-
violencia_domestica	Int	-
violencia_institucional	Int	-
violencia_laboral	Int	-
violencia_contra_libertad_reproductiva	Int	-
violencia_obstetrica	Int	-
violencia_mediatica	Int	-
violencia_otras	Int	-
vinculo_con_la_persona_agresora	varchar (50)	FK
genero_persona_agresora	varchar (50)	FK
poblacion_provincia	Int	-

Tabla Genero Agresor:

Contiene los datos del genero del agresor.

Atributos:

Tabla Genero Agresor		
Campo	Tipo de campo	Tipo de clave
ID_G_Agresor	varchar (50)	PK
genero de la persona agresora	varchar (50)	-

Tabla Vinculo

Detalla el vínculo que tiene el agresor con la victima.

Atributos:

Tabla Vinculo		
Campo	Tipo de campo	Tipo de clave
ID_vinculo	varchar (50)	PK
vínculo con la persona agresora	varchar (50)	-

Tabla País Nacimiento

En esta tabla se enumeran los países de nacimiento de las víctimas de violencia.

Atributos:

Tabla País Nacimiento		
Campo	Tipo de	Tipo de clave
	campo	clave
ID_Pais	Int	PK
pais nacimiento Victima	varchar (50)	-

Tabla Genero

Se mencionan los géneros posibles de la víctima.

Atributos:

Tabla Genero		
Campo	Tipo de campo	Tipo de clave
ID_Genero	varchar (50)	PK
genero Victima	varchar (50)	-

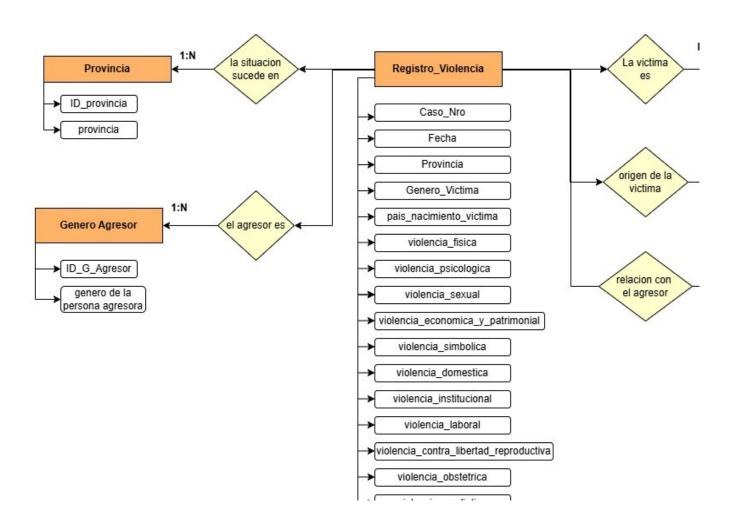
Tabla Provincia

Se especifican las provincias en las que ocurrieron los casos de violencia.

Atributos:

Tabla Provincia		
Campo	Tipo de campo	Tipo de clave
ID_provincia	varchar (50)	PK
provincia	varchar (50)	-

Modelo de entidad relación



Proceso de transformación de datos

Durante la etapa de preparación, se importaron las tablas desde archivos Excel a Power BI y se transformaron con Power Query. Las modificaciones incluyeron:

- Corrección de tipos de datos (por ejemplo, convertir fechas al formato correcto y los campos números en Int).
- Estandarización de textos (minúsculas, eliminación de espacios innecesarios).
- Eliminación de registros incompletos o inválidos.
- Creación de una tabla adicional 'Tipos Violencia' mediante DAX para simplificar el cálculo individual de cada tipo de violencia y usarlo en visualizaciones.
- Adición de una tabla de calendario con jerarquía de fechas (año, mes, semestre) para análisis temporal y su unión con el modelo entidad relación.
- Consolidación de campos de violencia en una medida dinámica para representar casos positivos.

El modelo se estructuró de forma relacional, uniendo las tablas por medio de claves primarias y foráneas que se definieron al momento relacionar cada una de las entidades.

La fórmula en DAX para la tabla creada fue la siguiente:

```
TiposViolencia =

DATATABLE(

"Tipo", STRING,

{

    {"Violencia Física"},
    {"Violencia Psicológica"},
    {"Violencia Económica"},
    {"Violencia Sexual"},
    {"Violencia Simbólica"},
    {"Violencia Institucional"},
    {"Violencia Obstétrica"},
    {"Violencia Laboral"},
    {"Violencia Doméstica"},
    {"Violencia contra la libertad reproductiva"},
    {"Violencia otras"},
    {"Violencia mediática"}
}

}
```

Esto es para poder usar la suma o cuenta de casos positivos de cada uno de los tipos de violencia como medida para generar algunos gráficos de interés general.

Se creó la tabla calendario siguiendo la formula dax que entrego el profesor y además de esto se hizo también la tabla de medidas para poder anidar allí las necesarias en cada uno de los gráficos y tarjetas.

Las medidas creadas fueron:

- <u>Cantidad total por género</u>: medida calculada para la tarjeta correspondiente:

- Cantidad total por género = COUNT('Registro Violencia'[ID Genero])

- <u>Count casos</u>: para poder contabilidad de manera segmentada la cantidad de casos dependiendo del gráfico:

count_casos = COUNT(Registro_Violencia[Caso_Nro])

Edad promedio víctima: Se hizo para usar en las tarjetas:

Edad Promedio Victima = AVERAGE(Registro_Violencia[edad_Victima])

<u>Fecha/hora</u>: Para mantener los datos de fecha actualizados en la caratula:

Fecha/hora = NOW()

<u>TotalporTipoViolencia</u>: Es una dax que permite cuantificar el número de casos por cada uno de los tipos de violencia y agruparlos de esa forma:

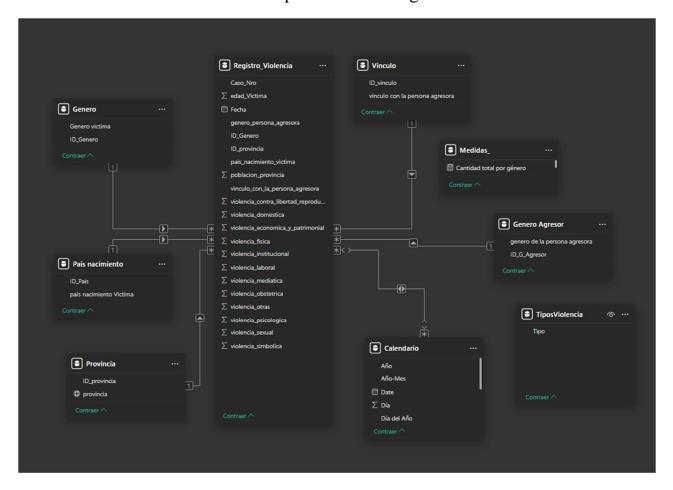
```
TotalPorTipoViolencia =
SWITCH(
  SELECTEDVALUE(TiposViolencia[Tipo]),
  "Violencia Física", SUM(Registro Violencia [violencia física]),
  "Violencia Psicológica",
SUM(Registro Violencia[violencia psicologica]),
  "Violencia Económica",
SUM(Registro Violencia[violencia economica y patrimonial]),
  "Violencia Sexual", SUM(Registro Violencia[violencia sexual]),
  "Violencia Simbólica", SUM(Registro Violencia[violencia simbolica]),
  "Violencia Institucional",
SUM(Registro Violencia[violencia institucional]),
  "Violencia Obstétrica", SUM(Registro Violencia[violencia obstetrica]),
  "Violencia Laboral", SUM(Registro Violencia[violencia laboral]),
  "Violencia Doméstica".
SUM(Registro_Violencia[violencia domestica]),
  "Violencia contra la libertad reproductiva",
SUM(Registro Violencia[violencia contra libertad reproductiva]),
  "Violencia otras", SUM(Registro Violencia[violencia otras]),
  "Violencia mediática", SUM(Registro Violencia[violencia mediatica])
```

Se creó también la tabla calendario con dax:

```
Calendario =
ADDCOLUMNS(
    CALENDAR(DATE(2010,1,1), DATE(2030,12,31)),
    "Año", YEAR([Date]),
    "Mes", FORMAT([Date], "MMMM"),
    "Mes Número", MONTH([Date]),
    "Año-Mes", FORMAT([Date], "YYYY-MM"),
    "Trimestre", "Q" & FORMAT([Date], "Q"),
    "Semana", WEEKNUM([Date], 2),
    "Día", DAY([Date]),
    "Nombre Día", FORMAT([Date], "dddd"),
    "Día del Año", FORMAT([Date], "DDD")
)
```

Y la tabla de medidas con una columna vacía que luego se eliminó para dejar anidadas allí todas las medidas calculadas creadas.

Finalmente, una vez creadas las tablas necesarias el modelo entidad relación se actualizo en Power Bi quedando de la siguiente manera:



La tabla Calendario fue unida al modelo para crear la jerarquía de fechas y poder ser utilizada en el análisis histórico para generar gráficos de interés sobre la evolución de los casos a lo largo del tiempo.

Detalle de Solapas

El análisis esta segmentado en cuatro solapas con una caratula de presentación:

Caratula



La primera Solapa es la portada en donde podemos ver el titulo introductorio del trabajo y en el cual se incluyeron 4 botones de navegación que se configuraron para llevar a cada una de las solapas en las que anidan los distintos tipos de análisis que se decidieron presentar para el trabajo. Además de esto se incluyó la formula *now()* para mostrar la última actualización de la presentación y actualización de datos.

Glosario

Glosario



Violencia de Genero: Toda conducta basada en el género que cause daño físico, sexual, psicológico, económico o simbólico a una persona, especialmente cuando se ejerce contra mujeres o identidades feminizadas.

violencia física: Acciones que implican el uso de la fuerza sobre el cuerpo de otra persona, provocando dolor, lesiones o riesgo para la vida.

violencia psicológica: Conductas que dañan la salud emocional, la autoestima o la estabilidad mental de una persona, a través de humillaciones, amenazas, manipulación o aislamiento.

violencia sexual: Cualquier acción que implique una invasión o imposición sexual sin consentimiento, incluyendo abuso, acoso y violación.

violencia económica y patrimonial: Control, limitación o privación de recursos económicos o bienes personales, con el fin de someter o dominar a la otra persona.

violencia simbólica: Reproducción de desigualdades de género a través de mensajes, imágenes, estereotipos o prácticas culturales que perpetúan la subordinación de las mujeres y disidencias.

violencia domestica: Violencia ejercida en el ámbito del hogar o de relaciones afectivas, por parte de personas con las que se mantiene o mantuvo un vínculo familiar o convivencial.

violencia institucional: Acciones u omisiones de organismos públicos que obstaculizan, dilatan o impiden el acceso a derechos de manera arbitraria o discriminatoria.

violencia laboral: Maltrato, acoso o discriminación en el ámbito del trabajo que afecta la dignidad, la integridad o el acceso a oportunidades laborales.

violencia contra libertad reproductiva: Imposición, negación o interferencia sobre las decisiones vinculadas a la reproducción, como el acceso a métodos anticonceptivos, abortos legales o maternidad/paternidad deseada.

violencia obstétrica: Prácticas deshumanizadas hacia personas gestantes durante el embarazo, parto o posparto, que incluyan maltrato, intervenciones innecesarias o falta de consentimiento informado.

violencia mediática: Difusión de contenidos que refuercen estereotipos de género, cosifiquen cuerpos o promuevan discriminación, mediante medios de comunicación masiva.

Genero No binario: Identidad de género que no se ajusta exclusivamente a las categorías tradicionales de hombre o mujer. Puede incluir identidades fluidas, mixtas o fuera del binarismo.

Transgénero: Persona cuya identidad de género no coincide con el sexo asignado al nacer. Puede o no modificar su cuerpo o nombre legal.

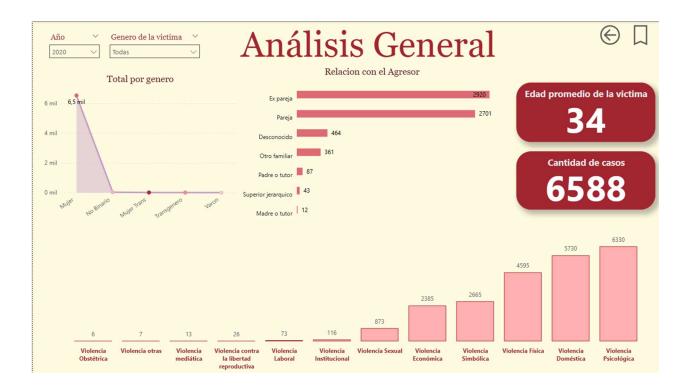
Travesti: Identidad de género presente especialmente en América Latina, que reivindica una vivencia feminizada no hegemónica, disidente del binarismo, y suele implicar una expresión corporal, política y cultural propia.

La segunda solapa contiene las definiciones de los conceptos mas importantes del dataset para que cualquier persona pueda entender de lo que se está hablando y entender cada una de las dimensiones que aparecen en los gráficos y tarjetas a lo largo del dashboard.

Funcionalidades interactivas:

• **Botón de navegación**: ofrece la posibilidad de volver a la portada desde cualquier punto del informe para acceder a otras solapas de análisis.

Análisis General



Esta primera solapa permite observar y analizar los casos a nivel global, brindando una visión panorámica del contenido del dataset. Sirve como punto de partida para comprender las principales características de las denuncias registradas y las variables más relevantes vinculadas a los distintos tipos de violencia, el perfil de las víctimas y la relación con el agresor

Elementos visuales incluidos:

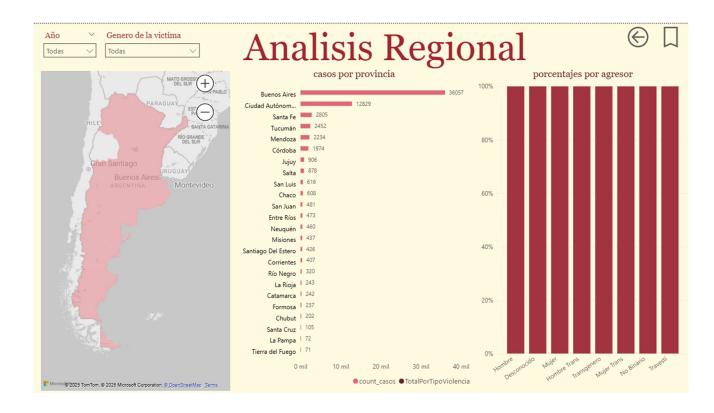
- Gráfico de columnas "Relación con el agresor": muestra la cantidad de casos según el vínculo de la víctima con el agresor. Se destacan las categorías "Ex pareja" y "Pareja", que concentran la mayor cantidad de registros, seguidas por "Desconocido" y "Otro familiar". Esta visualización permite identificar rápidamente los vínculos más frecuentes en las situaciones de violencia denunciadas.
- Gráfico de columnas apiladas "Tipos de violencia": se presenta la distribución por tipo de violencia ejercida. Las más recurrentes son la Violencia Psicológica, la Violencia Doméstica y la Violencia Física, lo que permite dimensionar qué formas de violencia están más presentes en los registros.

- Tarjeta con edad promedio: indica que la edad promedio de las víctimas es de 34 años, lo que permite situar el rango etario más afectado.
- Tarjeta con cantidad total de casos: muestra una cifra total aproximada de denuncias, lo que ayuda a dimensionar el alcance del fenómeno.

Funcionalidades interactivas:

- Segmentadores por año y género de la víctima: permiten filtrar todos los gráficos y tarjetas de la solapa para realizar análisis más focalizados según el período temporal o el género. Esto facilita observar variaciones o patrones específicos dentro de los datos.
- **Botón de navegación**: ofrece la posibilidad de volver a la portada desde cualquier punto del informe para acceder a otras solapas de análisis.
- **Botón de reinicio**: permite borrar todos los filtros aplicados y volver a la vista inicial con los datos completos, útil para reiniciar el análisis sin reiniciar todo el archivo.

Análisis Regional



Esta solapa se centra en la **distribución geográfica** de los casos de violencia, permitiendo identificar qué provincias registran mayor cantidad de denuncias y cómo se distribuyen los distintos tipos de agresores en cada región.

El objetivo principal es observar las localizaciones más afectadas y comprender el fenómeno en función del territorio, aportando así una dimensión regional al análisis.

Elementos visuales incluidos:

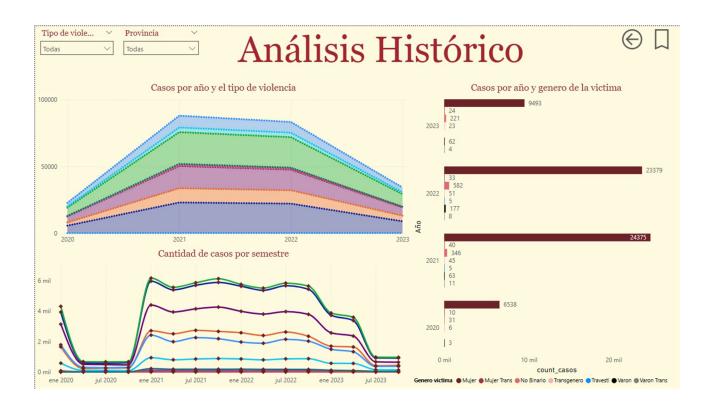
• Mapa de calor interactivo: muestra la densidad de casos por provincia en un mapa de Argentina. Las tonalidades más intensas indican mayor cantidad de denuncias, facilitando la visualización espacial del fenómeno.

- Gráfico de barras "Casos por provincia": presenta de forma ordenada la cantidad total de denuncias registradas en cada jurisdicción. Se destaca la provincia de Buenos Aires como la más afectada, seguida por Ciudad Autónoma de Buenos Aires, Santa Fe y Tucumán.
- Gráfico de barras apiladas "Porcentaje por agresor": permite observar cómo se distribuyen los diferentes tipos de agresores (ex pareja, pareja, desconocido, entre otros) en términos porcentuales. Aunque las proporciones se mantienen relativamente estables entre provincias, esta visualización permite detectar matices en la composición del vínculo con el agresor según la región.

Funcionalidades interactivas:

- Segmentadores por año y género de la víctima: permiten filtrar los gráficos y el mapa para realizar análisis focalizados según el período o el género, lo que mejora la capacidad de exploración y análisis por parte del usuario.
- Botón de navegación: posibilita regresar fácilmente a la portada del informe para continuar con otras secciones del análisis.
- **Botón de reinicio**: permite borrar todos los filtros aplicados y volver a la vista inicial con los datos completos, útil para reiniciar el análisis sin reiniciar todo el archivo.

Análisis Histórico



Esta sección tiene como propósito observar la **evolución temporal** de los casos de violencia reportados entre los años 2020 y 2023, diferenciando tanto por tipo de violencia como por género de la víctima.

El análisis histórico permite identificar **tendencias**, **picos específicos** y **disminuciones** en los registros, ofreciendo una visión global del comportamiento de los casos a lo largo del tiempo.

Elementos visuales incluidos:

- Gráfico de área "Casos por año y tipo de violencia": permite visualizar cómo varió la cantidad de casos según el tipo de violencia en cada año. Se observa un fuerte aumento entre 2020 y 2021, un leve descenso en 2022 y una baja más pronunciada en 2023.
- Gráfico de líneas "Cantidad de casos por semestre": muestra la evolución semestral, permitiendo identificar cambios más finos y estacionales. La visualización resalta que los picos de denuncias se concentran en el primer semestre de cada año, con una leve baja hacia la segunda mitad.

• Gráfico de barras "Casos por año y género de la víctima": presenta los registros agrupados por año y desglosados según género auto percibido. Se destacan las mujeres cis como el grupo más afectado en todos los períodos analizados, seguidas por varones y otras identidades de género en menor proporción.

Funcionalidades interactivas:

- Segmentadores por tipo de violencia y provincia: permiten filtrar el análisis para focalizar en tipos de violencia específicos o comparar entre provincias.
- **Botón de navegación**: permite volver fácilmente a la portada del informe.
- **Botón de reinicio**: permite borrar todos los filtros aplicados y volver a la vista inicial con los datos completos, útil para reiniciar el análisis sin reiniciar todo el archivo.

Conclusión

El presente análisis, basado en registros oficiales y normalizados, pone en evidencia patrones estructurales de la violencia de género en Argentina. Los datos no solo muestran que la mayoría de las víctimas son mujeres cis, sino que revelan que el tipo de vínculo con el agresor es en su mayoría íntimo —pareja o ex pareja— y que las formas más comunes de agresión son la violencia psicológica y doméstica. La edad promedio de las víctimas, alrededor de los 34 años, indica que estas situaciones afectan especialmente a mujeres en plena etapa laboral, social y reproductiva.

Estos hallazgos permiten trazar líneas de acción claras y urgentes. En primer lugar, es imperativo fortalecer las políticas de prevención centradas en vínculos afectivos violentos, desde la educación sexual integral hasta programas de concientización en medios, escuelas, clubes y espacios laborales. Es necesario trabajar no solo con las víctimas, sino con potenciales agresores, apuntando a la deconstrucción de modelos de masculinidad que reproducen control, celos, dominio y violencia.

En segundo lugar, los datos revelan focos regionales claros —Buenos Aires, CABA, Santa Fe y Tucumán— lo que permite **descentralizar los recursos y crear centros de atención integral en zonas priorizadas**, con equipos interdisciplinarios capacitados en contención, asesoría legal, asistencia psicológica y alojamiento inmediato.

Tercero, la franja etaria más afectada requiere **políticas económicas y laborales con enfoque de género**, que garanticen autonomía económica para que las mujeres puedan salir de círculos de violencia sin poner en riesgo su subsistencia o la de sus hijos. Esto implica fortalecer el acceso a empleo formal, redes de cuidado, subsidios de emergencia y vivienda segura.

En síntesis, los datos no solo reflejan la violencia: la hacen visible, trazan rutas posibles y permiten planificar intervenciones específicas. La violencia de género no es aleatoria ni inevitable; es prevenible si se actúa con decisión, estrategia y sensibilidad social. Este trabajo apunta a contribuir en esa dirección.