

PROYECTO FINAL - LINEA 144

INDICE

PROYECTO FINAL - LINEA 144.....	1
INFORME EJECUTIVO.....	2
Introducción:.....	2
Descripción de la temática de los datos:.....	3
Alcance del Proyecto:.....	3
Objetivo del Proyecto:.....	4
Nivel de Aplicación (Táctico):.....	4
Conclusiones:.....	5
INFORME TECNICO.....	9
Introducción:.....	9
Detalles del Dataset:.....	10
Normalización de Datos en SQL:.....	12
Explicación del Diagrama Obtenido en SQL:.....	14
Organización de Datos en POWER BI:.....	20
Explicación del Diagrama Obtenido en POWER BI:.....	29
Herramientas Utilizadas:.....	40
ANEXO.....	41
CARPETA DRIVE CON DICHOS ARCHIVOS.....	41

INFORME EJECUTIVO

Introducción:

- **Análisis de la Línea 144**
 - Este informe presenta un análisis basado en los datos poseídos sobre la Línea 144, un servicio de atención a víctimas de violencia de género perteneciente al Ministerio de Justicia, publicados por el **Portal Nacional de Datos Públicos** (datos.gob.ar).
El análisis abarca un total de ochenta y nueve mil quinientas cuarenta (89,540) víctimas, y abarca un período comprendido entre el 1 de enero de 2020 y el 30 de junio de 2023.
El objetivo de este análisis es determinar los tipos de violencia más denunciados, identificar las modalidades en que ocurren y analizar su frecuencia por trimestre.

Descripción de la temática de los datos:

- Base de datos Línea 144 Argentina
 - Este dataset¹ contiene información correspondiente a **comunicaciones recibidas por la Línea 144** que abarcan desde el año 2020 hasta mediados del 2023; dicha línea telefónica brinda **atención a personas en situación de violencia de género**.

Alcance del Proyecto:

1. **Cantidad de Registros:** El dataset inicial abarca un total de ochentinueve mil quinientos cuarenta(89540) registros/filas y veintiún(21) columnas.
2. **Variables Disponibles:** Algunas de las variables más importantes utilizadas para el análisis son:
 - a. La fecha de registro de cada llamada, para determinar la frecuencia de los diferentes tipos de violencia y modalidades.
 - b. Los diferentes tipos de violencia que declaró sufrir la víctima (física, psicológica, económica/patrimonial, etc.).
 - c. Bajo qué modalidad/es ocurrieron (institucional, laboral, etc.).
3. **Rango Temporal:** El dataset abarca un período comprendido entre el **1 de enero de 2020** y el **30 de junio de 2023**.
4. **Análisis a Realizar:** Se efectuará un recuento de los casos partiendo desde los diferentes tipos de violencia y a su vez desde las diferentes modalidades obteniendo así los diferentes porcentajes de ocurrencia. Posteriormente se analizarán en conjunto con las fechas para determinar su frecuencia;
5. **Limitaciones:** Debido a que ochenta y cuatro mil quinientas veintisiete(84527) víctimas no registraron modalidad alguna, no se ha podido realizar ningún análisis en cuanto a la relación que tenía con la variable principal (tipos de violencia).
Además se han observado registros vacíos o valores nulos en algunas variables, sin embargo estos últimos se han normalizados respectivamente, los mismos no influyen con el objetivo del proyecto.
 - a. Más información en el [Informe Técnico](#).

¹ **Dataset:** conjunto de datos almacenados que poseen relación entre sí y que, generalmente, se encuentran en forma de tablas (filas y columnas).

Objetivo del Proyecto:

- Base de datos Línea 144 Argentina
 - El objetivo es comprender **cuál o cuáles** son los **tipos de violencia más registrados, bajo qué modalidades ocurren y la frecuencia de las mismas por trimestre**.
Permitiendo así crear **medidas efectivas y específicas** frente a los **tipos y modalidades de violencia más comunes y frecuentes**.

Nivel de Aplicación (Táctico):

- Usuario objetivo
 - Dicho análisis está dirigido a los sectores del Ministerio de Justicia responsables de la implementación y gestión de políticas a mediano plazo, a fin de que se tomen decisiones operativas que ayuden con la creación de medidas específicas en función de los resultados obtenidos.

Conclusiones:

- Los *tipos de violencias que más se reportaron* son la violencia **psicológica** y **doméstica**.
 - **Psicológica**: con un total de ochenticuatro mil novecientos noventa y dos(84992) equivalente al 94,92% del dataset.
 - **Doméstica**: con un total de ochenta y un mil trescientos treinta y cuatro(81334) equivalente al 90,84% del dataset.
 - **Demás violencias**:
 - i. **Física**: con un total de cincuenta y ocho mil novecientos sesenta y cuatro(58964) equivalente al 65,85 % del dataset.
 - ii. **Económica Patrimonial**: con un total de treinta y cinco mil novecientos treinta y uno(35931) equivalente al 40,13 % del dataset.
 - iii. **Simbólica**: con un total de treinta mil quinientos ochenta y dos(30582) equivalente al 34,15 % del dataset.
 - iv. **Sexual**: con un total de doce mil doscientos treinta y nueve(12239) equivalente al 13,67 % del dataset.
- La combinación de violencia número “34” y número “2”, son las más comunes.
 - Combinación número “34”: Las víctimas que registran esta combinación sufren violencia psicológica, doméstica y física; la cantidad de víctimas es de diez y nueve mil cuatrocientas cuarenta y tres(19443).
 - Combinación número “2”: Las víctimas que registran esta combinación sufren violencia psicológica, doméstica, física y económica patrimonial; la cantidad de víctimas es de doce mil novecientas setenta(12970).
 - **DETALLE**:
 - i. Téngase en cuenta que el total de víctimas del dataset es de ochenta y nueve mil quinientos cuarenta(89540).
 - ii. Téngase en cuenta que una víctima podría sufrir varios tipos de violencia.
- MODALIDADES
 - Conclusiones a partir de datos reales:
 - i. No se ha podido analizar con certeza bajo qué modalidades ocurrieron los diferentes tipos de violencias debido a que ochenta y cuatro mil quinientos veintisiete(84527) víctimas no registraron modalidad alguna.
 - 1. Solamente un total de cinco mil trece(5013) víctimas denunciaron sufrir violencia bajo una o más modalidades repartidas en:

- a. **Otras:** un total de dos mil ciento siete(2107) víctimas.
- b. **Institucional:** un total de mil trescientas cuatro(1304) víctimas.
- c. **Laboral:** un total de mil doscientas cincuenta y un(1251) víctimas.
- d. **Libertad Reproductiva:** un total de cuatrocientas(400) víctimas.
- e. **Obstétrica:** un total de ochenta y cinco(85) víctimas.
- f. **Mediática:** un total de cincuenta y tres(53) víctimas.

ii. **DETALLE:**

- 1. Téngase en cuenta que el total de víctimas del dataset es de ochenta y nueve mil quinientos cuarenta(89540).
- 2. Téngase en cuenta que una víctima podría sufrir violencia bajo varios tipos de modalidades.
- 3. Téngase en cuenta que los datos podrían haber sido mal exportados, por lo que sería recomendable realizar una prueba de integridad a la base de datos correspondiente para verificar y/o revisar o mejorar el procedimiento de ingreso de datos para así asegurar un buen estado de los mismos.

○ Conclusiones a partir de los datos falsos imputados:

- i. La *modalidad más reportada* fue la **Institucional**.
 - 1. Solamente quinientas(500) víctimas no registraron modalidad alguna.
 - a. **Institucional:** un total de cuarenta mil quinientas(40500) víctimas, equivalente al 45.43% del dataset.
 - b. **Obstétrica:** un total de veintinueve mil novecientas(29900) víctimas, equivalente al 33.39% del dataset.
 - c. **Laboral:** un total de 29,000 víctimas, equivalente al 32.38% del dataset.
 - d. **Libertad Reproductiva:** un total de 24,500 víctimas, equivalente al 27.36% del dataset.
 - e. **Otras:** un total de 21,000 víctimas, equivalente al 23.45% del dataset.
 - f. **Mediática:** un total de 19,000 víctimas, equivalente al 21.21% del dataset.

ii. **DETALLE:**

1. Téngase en cuenta que el total de víctimas del dataset es de ochenta y nueve mil quinientos cuarenta(89540).
 2. Téngase en cuenta que una víctima podría sufrir violencia bajo varios tipos de modalidades.
 3. **Téngase MUY en cuenta que estas conclusiones fueron hechas bajo datos falsos.**
- La combinación de modalidad número “5” lidera las estadísticas con exactamente siete mil(7000) víctimas; el resto de las combinaciones rondan entre cinco mil(5000) y mil(1000) víctimas.
 - i. La combinación número “5” establece que la víctima sufrió violencia bajo la modalidad “Institucional” y “Obstétrica”.
 - ii. **DETALLE:**
 1. **Téngase MUY en cuenta que estas conclusiones fueron hechas bajo datos falsos.**
- El promedio de llamadas a la Línea 144 es de un total de seis mil trescientos noventa y seis(6396) por trimestre.

En cuanto a los diferentes tipos de violencia y sus frecuencias:

 - Seis mil setenta y un(6071) víctimas por trimestre reportaron sufrir violencia psicológica.
 - Cinco mil ochocientos diez(5810) víctimas por trimestre reportaron sufrir violencia doméstica.
 - Cuatro mil doscientas doce(4212) víctimas por trimestre reportaron sufrir violencia física.
 - Dos mil quinientas sesenta y siete(2567) víctimas por trimestre reportaron sufrir violencia económica patrimonial.
 - Dos mil ciento ochenta y cuatro(2184) víctimas por trimestre reportaron sufrir violencia simbólica.
 - Ochocientas setenta y cuatro(874) víctimas por trimestre reportaron sufrir violencia sexual.
 - Los lunes, sábados y domingos son los días que más víctimas llaman a denunciar su caso, por ende sería conveniente revisar la calidad de la disponibilidad que ofrece la línea 144 en esos días.
 - Tomando solamente los años 2020, 2021 y 2022 los meses donde más casos de víctimas se reportan es en marzo y abril con ‘7.126’ y ‘7.047’ respectivamente; también agosto y mayo con ‘6958’ y ‘6928’.
 - (No se toma el año 2023 en cuenta ya que como abarca los dos primeros trimestres podría sesgar el recuento de víctimas por mes y por ende esta conclusión)

- Solamente doscientas noventa y un(291) víctimas no poseen registro alguno sobre una o más violencias sufridas, por lo que se recomendaría revisar, además de la integridad de los datos, el proceso con el cual se ingresan los mismos. También se recomendaría evaluar el proceso de selección con el cual han catalogado como “válidos” los datos en este dataset y evitar errores a futuro.
- Se ha observado que la provincia de Buenos Aires y la Ciudad Autónoma de Buenos Aires (CABA) son las más predominantes en el dataset (ocupando el 73,66% de los casos) por lo que podría sugerirse realizar propagandas en el resto de las provincias a fin de ampliar el uso y alcance de la línea 144.
- Se sugerirá crear una nueva opción dentro de la categoría “Relación con el Agresor” el cual se denomine “Tutor” para de esta forma evitar redundancias del estilo “Madre o Tutor”, “Padre o Tutor”, etc.

INFORME TECNICO

Introducción:

- **Proceso madurativo del proyecto**
 - En este informe se detalla el proceso completo de desarrollo del proyecto, desde la recopilación inicial de datos hasta el análisis final realizado en Power BI. A lo largo del documento, se explican las decisiones tomadas y los pasos realizados en cada fase del proyecto, con el objetivo de proporcionar una visión más amplia y técnica de todo el trabajo realizado.

Detalles del Dataset:

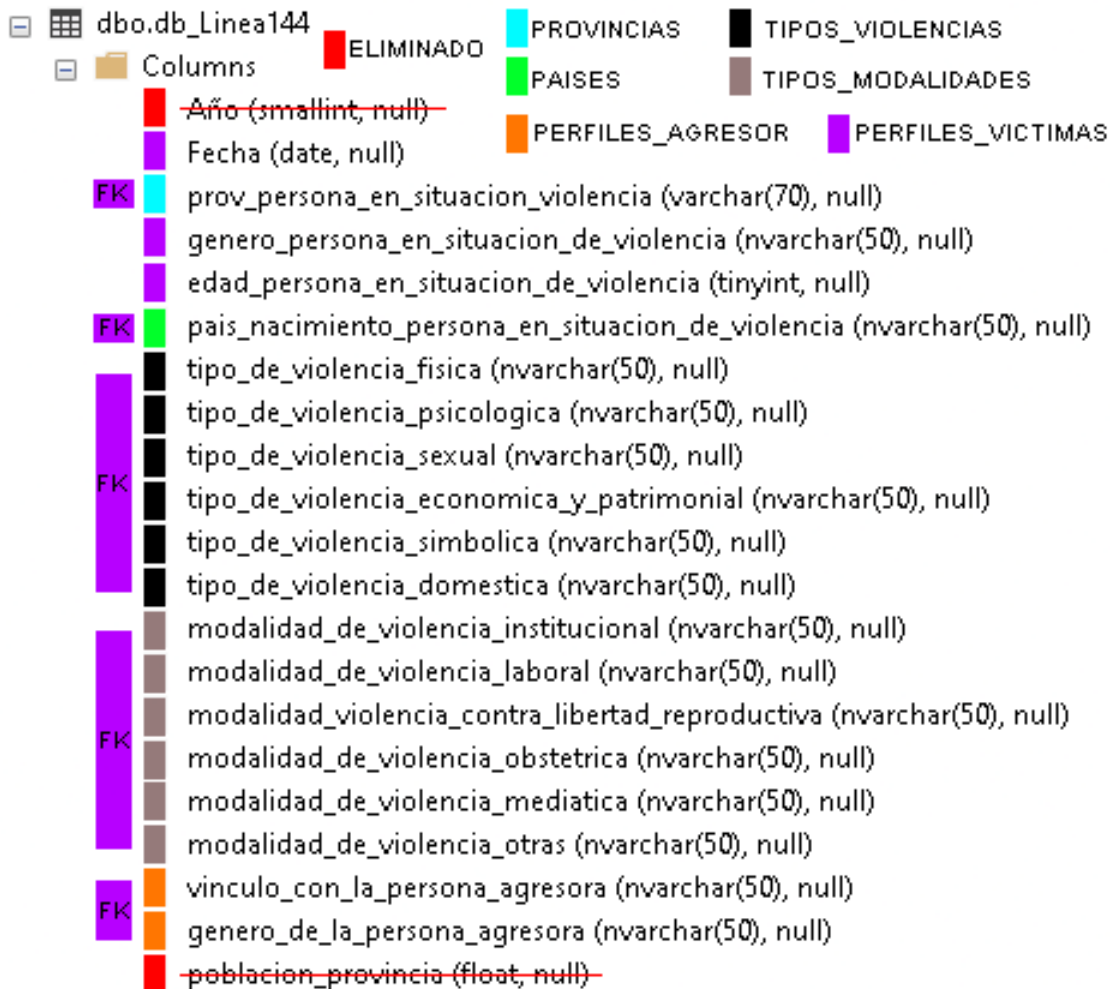
- [ARCHIVO DATASET INICIAL](#)
 - Archivo.csv
- [ARCHIVO DATASET SQL](#)
 - Archivo.bacpac (SQL SERVER)
- [ARCHIVO POWER BI](#)
 - Archivo.pbix (POWER BI)
- **VISUALIZACION DE COLUMNAS DEL DATASET INICIAL**

dbo.db_Lineas144

Columns

- Año (smallint, null)
- Fecha (date, null)
- prov_persona_en_situacion_violencia (varchar(70), null)
- genero_persona_en_situacion_de_violencia (nvarchar(50), null)
- edad_persona_en_situacion_de_violencia (tinyint, null)
- pais_nacimiento_persona_en_situacion_de_violencia (nvarchar(50), null)
- tipo_de_violencia_fisica (nvarchar(50), null)
- tipo_de_violencia_psicologica (nvarchar(50), null)
- tipo_de_violencia_sexual (nvarchar(50), null)
- tipo_de_violencia_economica_y_patrimonial (nvarchar(50), null)
- tipo_de_violencia_simbolica (nvarchar(50), null)
- tipo_de_violencia_domestica (nvarchar(50), null)
- modalidad_de_violencia_institucional (nvarchar(50), null)
- modalidad_de_violencia_laboral (nvarchar(50), null)
- modalidad_violencia_contra_libertad_reproductiva (nvarchar(50), null)
- modalidad_de_violencia_obstetrica (nvarchar(50), null)
- modalidad_de_violencia_mediatica (nvarchar(50), null)
- modalidad_de_violencia_otras (nvarchar(50), null)
- vinculo_con_la_persona_agresora (nvarchar(50), null)
- genero_de_la_persona_agresora (nvarchar(50), null)
- poblacion_provincia (float, null)

• SELECCION DE LAS COLUMNAS PARA TABLAS DEL DIAGRAMA



• EXISTENCIA DE NULOS POR COLUMNA

- "pais_nacimiento_persona_en_situacion_de_violencia" = 30281
- "edad_persona_en_situacion_de_violencia" = 17278
- "genero_de_la_persona_agresora" = 9213
- "vinculo_con_la_persona_agresora" = 3395
- "genero_persona_en_situacion_de_violencia" = 1904
- "poblacion_provincia" = 1312

Normalización de Datos en SQL:

- CAMBIOS REALIZADOS CON RESPECTO AL DATASET INICIAL

- Se han removido las columnas:
 - “Año” por falta de utilidad ya que la columna “Fecha” cumpliría exactamente la misma función.
 - “poblacion_provincia” por falta de utilidad ya que la cantidad de datos es relativamente poca como para realizar un análisis con dicha variable.
- Se ha modificado la columna:
 - “pais_nacimiento_persona_en_situacion_de_violencia” debido al uso de gentilicios como por ejemplo “Española”, “Francesa”, “Italiana”, etc. y se ha colocado como reemplazo, el nombre del país al cual hace referencia dicho gentilicio, es decir, “España”, “Francia”, “Italia”, etc.
- Se aplicó una técnica de One Hot Encoding para transformar los datos “Si” y “No”, de las columnas “tipo_de_violencia_X” y “modalidad_de_violencia_X” en valores de 0 (para “No”) y 1 (para “Si”); De esta manera se podrán realizar diferentes análisis numéricos y estadísticos sobre columnas que inicialmente eran categóricas.

- MANEJO DE NULOS:

- Variables Categóricas
 - “pais_nacimiento_persona_en_situacion_de_violencia”
 - “genero_de_la_persona_agresora”
 - “vinculo_con_la_persona_agresora”
 - “genero_persona_en_situacion_de_violencia”
 - “prov_persona_en_situacion_de_violencia”
 - Para las variables categóricas se decidió renombrar todos los registros “NULL” por “Desconocido”; en el caso de la columna “prov_persona_en_situacion_de_violencia” también se cambió su forma inicial “Sin datos”, por “Desconocido”.
- Variables Numéricas
 - “edad_persona_en_situacion_de_violencia”
 - Para la única variable numérica se observó que posee un total de diecisiete mil doscientos setenta y ocho(17278) registros nulos, lo que equivale al 19,2% del dataset, por lo tanto se decidió rellenar dichos datos con el promedio de la edad del dataset, es decir, treinticinco(35) ya que



esta acción no ocasiona un sesgo significativo.

Explicación del Diagrama Obtenido en SQL:

- DESCRIPCION GENERAL

- El diagrama contiene seis(6) tablas en total, de las cuales cinco(5) son tablas dimensionales, que poseen información característica respectivamente, y una(1) es una tabla de hechos, la cual une y centraliza a todas ellas, formando así un **esquema en estrella**.

- DESCRIPCION DE TABLAS

- PAISES

- Tabla dimensional que contiene de manera única los distintos países existentes en el dataset; contiene cuarenta y dos(42) posibilidades.

- paisID: Llave primaria (PK) para identificar de manera única cada país, columna tipo (INT).
- nombre_pais: Nombre del país, columna tipo (NVARCHAR).

	Results	Messages
	paisID	nombre_pais
1	1	Alemania
2	2	Arabia
3	3	Argentina
4	4	Armenia
5	5	Australia
6	6	Bélgica
7	7	Bolivia
8	8	Brasil
9	9	Canadá
10	10	Chile
11	11	Colombia
12	12	Costa Rica
13	13	Cuba
14	14	Desconocido
15	15	Ecuador
16	16	El Salvador

- PROVINCIAS

- Tabla dimensional que contiene de manera única las distintas provincias existentes en el dataset junto con la población de cada una registrada hasta entonces; contiene veinticinco(25) posibilidades.

- provincialID: Llave primaria (PK) para identificar de manera única cada provincia, columna tipo (INT).
- nombre_provincia: Nombre de la provincia, columna tipo (NVARCHAR).

	Results	Messages
	provincialD	nombre_provincia
1	1	Ciudad Autónoma de Buenos Aires
2	2	San Luis
3	3	San Juan
4	4	Buenos Aires
5	5	Misiones
6	6	Córdoba
7	7	Salta
8	8	Santiago Del Estero
9	9	Tucumán
10	10	Jujuy
11	11	Desconocido
12	12	La Pampa
13	13	La Rioja
14	14	Catamarca
15	15	Neuquén
16	16	Santa Fe

○ TIPOS_VIOLENCIA

- Tabla dimensional que posee los diferentes tipos de violencia existentes en el dataset; contiene sesenta y tres (63) combinaciones.
 - tipoviolencialD: Llave primaria (PK) para identificar de manera única cada tipo de violencia, columna tipo (INT).
 - fisica: Indicador binario de violencia física, columna tipo (INT).
 - psicologica: Indicador binario de violencia psicológica, columna tipo (INT).
 - sexual: Indicador binario de violencia sexual, columna tipo (INT).
 - economica_patrimonial: Indicador binario de violencia económica y patrimonial, columna tipo (INT).
 - simbolica: Indicador binario de violencia simbólica, columna tipo (INT).
 - domestica: Indicador binario de violencia doméstica, columna tipo (INT).

	tipoviencialD	fisica	psicologica	sexual	economica_patrimonial	simbolica	domestica
1	1	0	0	0	0	0	0
2	2	1	1	0	1	0	1
3	3	1	0	1	1	1	1
4	4	1	0	0	1	1	1
5	5	1	0	1	1	0	0
6	6	0	0	1	1	1	0
7	7	1	1	1	0	1	1
8	8	0	0	1	1	0	1
9	9	1	0	0	0	0	1
10	10	0	0	0	1	0	1
11	11	0	1	1	1	1	1
12	12	1	0	1	0	1	0
13	13	1	0	1	0	0	1
14	14	1	0	0	1	0	0
15	15	1	1	1	1	0	1
16	16	0	1	0	0	1	0

TIPOS_MODALIDAD

- Tabla dimensional que posee los diferentes tipos de modalidades de violencia existentes en el dataset; contiene veintidós(22) combinaciones.
 - tipomodalidadID: Llave primaria (PK) para identificar de manera única cada tipo de modalidad de violencia; columna tipo (INT).
 - institucional: Indicador binario de violencia institucional, columna tipo (INT).
 - laboral: Indicador binario de violencia laboral, tipo (INT).
 - libertad_reproductiva: Indicador binario de violencia contra la libertad reproductiva, columna tipo (INT).
 - obstetrica: Indicador binario de violencia obstétrica, tipo (INT).
 - mediatica: Indicador binario de violencia mediática, columna tipo (INT).
 - otras: Indicador binario de otras modalidades de violencia, columna tipo (INT).

Results		Messages					
	tipomodalidadID	institucional	laboral	libertad_reproductiva	obstetrica	mediatica	otras
1	1	0	0	0	0	0	0
2	2	1	0	1	1	0	0
3	3	1	0	0	0	0	1
4	4	0	0	0	1	0	1
5	5	1	0	0	1	0	0
6	6	1	1	0	0	0	0
7	7	1	1	1	0	0	0
8	8	0	0	0	0	0	1
9	9	1	1	0	1	0	0
10	10	1	0	0	0	1	0
11	11	0	1	1	0	0	0
12	12	1	0	1	0	0	0
13	13	0	1	0	0	0	0
14	14	1	0	0	0	0	0
15	15	0	0	0	0	1	0
16	16	0	0	1	1	0	0

PERFILES_AGRESOR

- Tabla dimensional que contiene los diferentes tipos de perfiles de agresores existentes del dataset; contiene cuarenta y tres(43) combinaciones.
 - tipoagresorID: Llave primaria (PK) para identificar de manera única cada perfil de agresor, columna tipo (INT).
 - genero_agresor: Género del agresor, columnas tipo (NVARCHAR).
 - vinculo_agresor: Vínculo del agresor con la víctima, columna tipo (NVARCHAR).

Results		Messages	
	tipoagresorID	genero_agresor	vinculo_agresor
1	1	Varon Trans	Pareja
2	2	Varon	Superior jerarquico
3	3	Varon	Otro familiar
4	4	Varon	Ex pareja
5	5	Transgenero	Pareja
6	6	Mujer Trans	Ex pareja
7	7	Otro	Otro familiar
8	8	Otro	Superior jerarquico
9	9	Mujer	Madre o tutor
10	10	Desconocido	Otro
11	11	Otro	Ex pareja
12	12	Desconocido	Madre o tutor
13	13	Desconocido	Ex pareja
14	14	Desconocido	Padre o tutor
15	15	Varon Trans	Otro
16	16	Desconocido	Otro familiar

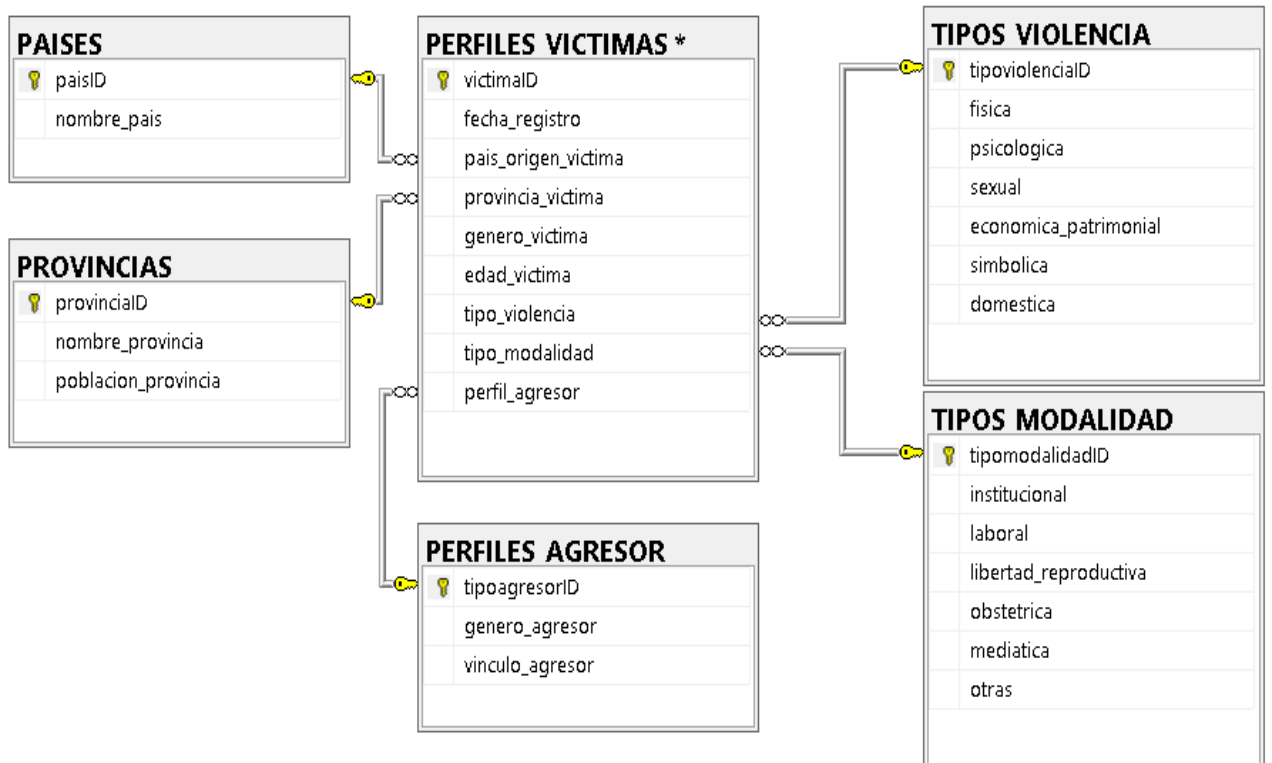
○ PERFILES_VICTIMAS | Tabla de Hechos

- Tabla de hechos que conecta las características de las diferentes entidades en el esquema de datos que están relacionadas con las víctimas y los incidentes de violencia de género; contiene ochentinueve mil quinientos cuarenta(89540) registros/filas.
 - victimaID: Llave primaria (PK) para identificar de manera única cada víctima, columna tipo (INT).
 - fecha_registro: Fecha en la que se registró el caso, columna tipo (DATE).
 - pais_origen_victima: Llave foránea (FK) de la tabla “PAISES” para identificar el país de origen de la víctima, columna tipo (INT).
 - provincia_victima: Llave foránea (FK) de la tabla “PROVINCIAS” para identificar la provincia de la víctima, columna tipo (INT).
 - genero_victima: Género de la víctima, columna tipo (NVARCHAR).
 - edad_victima: Edad de la víctima, columna tipo (INT).
 - tipo_violencia: Llave foránea (FK) de la tabla “TIPOS_VIOLENCIA” para identificar el tipo de violencia sufrida por la víctima, columna tipo (INT).

- tipo_modalidad: Llave foránea (FK) de la tabla “TIPOS_MODALIDAD” para identificar el tipo de modalidad sufrida por la víctima, columna tipo (INT).
- perfil_agresor: Llave foránea (FK) de la tabla “PERFILES_AGRESOR” para identificar el tipo de perfil del agresor, columna tipo (INT).

Results		Messages							
	victimID	fecha_registro	pais_origen_victima	provincia_victima	genero_victima	edad_victima	tipo_violencia	tipo_modalidad	perfil_agresor
1	1	2023-01-01	14	4	Mujer	35	34	1	26
2	2	2023-01-01	3	4	Mujer	19	25	1	28
3	3	2023-01-01	3	4	Mujer	46	26	1	26
4	4	2023-01-01	14	24	Mujer	28	33	1	26
5	5	2023-01-01	3	20	Mujer	34	56	1	4
6	6	2023-01-01	34	4	Mujer	38	2	1	4
7	7	2023-01-01	3	4	Mujer	57	2	1	26
8	8	2023-01-01	3	9	Mujer	35	26	1	26
9	9	2023-01-01	3	16	Mujer	39	42	1	4
10	10	2023-01-01	3	20	Mujer	24	15	1	26
11	11	2023-01-01	14	4	Mujer	35	34	1	4
12	12	2023-01-01	3	5	Mujer	22	2	1	26
13	13	2023-01-01	14	16	Mujer	52	34	1	26
14	14	2023-01-01	3	16	Mujer	65	2	1	4
15	15	2023-01-01	14	4	Mujer	26	2	1	26
16	16	2023-01-01	14	16	Mujer	41	2	1	26

• DIAGRAMA RESULTANTE EN SQL:



Organización de Datos en POWER BI:

- **CAMBIOS REALIZADOS CON RESPECTO AL ESQUEMA ESTRELLA OBTENIDO EN SQL**

- Implementación de tablas
 - Tabla “MEDIDAS”
 - Tabla que contiene veinte(22) medidas necesarias para la realización de los diferentes gráficos y por consiguiente su análisis.
 - Tabla “CALENDARIO”
 - Tabla creada a partir de la función “CALENDARAUTO()” que abarca un periodo desde la primera fecha encontrada en el dataset (1 de enero de 2020) hasta el 31 de diciembre del 2023 y está conectada con la tabla de hechos “PERFILES_VICTIMAS” a través de la columna “fecha_registro”.
 - Tabla “TablaCoordenadas”
 - Tabla importada a partir de un archivo “.csv” que contiene la latitud y longitud de todas las provincias del dataset; la misma se conecta a la tabla “PROVINCIAS” a través de “nombre_provincia”.
 - Tablas Desnormalizadas
 - “TablaViolenciasDesnormalizadas”: Tabla que almacena los diferentes nombres de los tipos de violencias, para cada tipo de combinación de violencia(ID) y que indica si posee o no ese tipo de violencia a través del número “0” para no y “1” para si.
 - “TablaModalidadesDesnormalizadas”: Tabla que almacena los diferentes nombres de los tipos de modalidades, para cada tipo de combinación de modalidad(ID) y que indica si posee o no ese tipo de modalidad a través del número “0” para no y “1” para si.
- Columnas agregadas
 - “IntervaloEdad” en “PERFILES_VICTIMAS”: Categoriza a las víctimas por intervalo en función de su edad.

```

1 IntervaloEdad =
2 SWITCH(
3     TRUE(),
4     [edad_victima] >= 1 && [edad_victima] <= 10, "1-10",
5     [edad_victima] > 10 && [edad_victima] <= 20, "11-20",
6     [edad_victima] > 20 && [edad_victima] <= 30, "21-30",
7     [edad_victima] > 30 && [edad_victima] <= 40, "31-40",
8     [edad_victima] > 40 && [edad_victima] <= 50, "41-50",
9     [edad_victima] > 50 && [edad_victima] <= 60, "51-60",
10    [edad_victima] > 60 && [edad_victima] <= 70, "61-70",
11    [edad_victima] > 70 && [edad_victima], "71-80+",
12    "Sin Datos"
13 )

```

- “tipo_modalidad_falsa” en “PERFILES_VICTIMAS”: Se imputan datos falsos ya que los originales impedían el análisis; esto es solo con fines académicos.

```

1 tipo_modalidad_falsa =
2 SWITCH(
3     TRUE(),
4     PERFILES_VICTIMAS[victimaID] > 89000, 23,
5     PERFILES_VICTIMAS[victimaID] > 88000, 22,
6     PERFILES_VICTIMAS[victimaID] > 85000, 21,
7     PERFILES_VICTIMAS[victimaID] > 80000, 20,
8     PERFILES_VICTIMAS[victimaID] > 75000, 19,
9     PERFILES_VICTIMAS[victimaID] > 70000, 18,
10    PERFILES_VICTIMAS[victimaID] > 65000, 17,
11    PERFILES_VICTIMAS[victimaID] > 60000, 16,
12    PERFILES_VICTIMAS[victimaID] > 55000, 15,
13    PERFILES_VICTIMAS[victimaID] > 50000, 14,
14    PERFILES_VICTIMAS[victimaID] > 45000, 13,
15    PERFILES_VICTIMAS[victimaID] > 40000, 12,
16    PERFILES_VICTIMAS[victimaID] > 35000, 11,
17    PERFILES_VICTIMAS[victimaID] > 30000, 10,
18    PERFILES_VICTIMAS[victimaID] > 25000, 9,
19    PERFILES_VICTIMAS[victimaID] > 20000, 8,
20    PERFILES_VICTIMAS[victimaID] > 18000, 7,
21    PERFILES_VICTIMAS[victimaID] > 17000, 6,
22    PERFILES_VICTIMAS[victimaID] > 10000, 5,
23    PERFILES_VICTIMAS[victimaID] > 5000, 4,
24    PERFILES_VICTIMAS[victimaID] > 2000, 3,
25    PERFILES_VICTIMAS[victimaID] > 500, 2,
26    PERFILES_VICTIMAS[victimaID] > 1, 1,
27    1)

```

● MEDIDAS

○ FALSAS

■ FALSO_TOTAL_VICTIMAS_SIN_MODALIDAD

- Esta medida cuenta las víctimas de la columna “tipo_modalidad_falsa” (la cual es falsa y fue creada con

fines académicos) que no poseen modalidad, es decir, que el tipo de modalidad sea igual a uno(1).

- ```

1 FALSO_TOTAL_VICTIMAS_SIN_MODALIDAD =
2 VAR sinmodalidad = CALCULATE([TOTAL_VICTIMAS], PERFILES_VICTIMAS[tipo_modalidad_falsa] = 1)
3 RETURN IF(sinmodalidad = 0, 0, sinmodalidad)

```

## ■ FALSO\_PORCENTAJE\_VICTIMAS\_POR\_TIPO\_MODALIDAD

- Esta medida es aplicada en el tooltip “tipos\_modalidades” (el cual es utilizado en el gráfico de barras de la solapa “MODALIDADES” denominado como “COMBINACION DE MODALIDADES”) y devuelve el porcentaje de víctimas por tipo de modalidad según el número de combinación.

- ```

1 FALSO_PORCENTAJE_VICTIMAS_POR_TIPO_MODALIDAD =
2 VAR ModalidadSeleccionada = SELECTEDVALUE(TablaModalidadesDesnormalizada[TipoModalidad])
3 VAR TotalVictimas = CALCULATE(COUNT(PERFILES_VICTIMAS[victimaID]), REMOVEFILTERS())
4 RETURN
5 SWITCH(
6     TRUE(),
7     ModalidadSeleccionada = "Institucional", DIVIDE(CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD
8     [Institucional] = 1), TotalVictimas),
9     ModalidadSeleccionada = "Laboral", DIVIDE(CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[laboral] = 1),
10    TotalVictimas),
11    ModalidadSeleccionada = "Libertad Reproductiva", DIVIDE(CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD
12    [libertad_reproductiva] = 1), TotalVictimas),
13    ModalidadSeleccionada = "Mediática", DIVIDE(CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[mediatica] =
14    1), TotalVictimas),
15    ModalidadSeleccionada = "Obstétrica", DIVIDE(CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[obstetrica] =
16    1), TotalVictimas),
17    ModalidadSeleccionada = "Otras", DIVIDE(CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[otras] = 1),
18    TotalVictimas),
19    BLANK()
20 )

```

■ FALSO_MODALIDAD_CON_MAS_VICTIMAS

- Esta medida falsa ayuda a determinar el tipo de modalidad con más víctimas en el dataset a partir de una variable que utiliza la función “MAXX()”, que toma el número máximo por cada variable, para luego compararlos obteniendo la modalidad correcta; Es implementada en la tarjeta de “INICIO”.

```

1 FALSO_MODALIDAD_CON_MAS_VICTIMAS =
2 VAR Institucional = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[institucional] = 1)
3 VAR Laboral = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[laboral] = 1)
4 VAR LibertadReproductiva = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[libertad_reproductiva] = 1)
5 VAR Mediática = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[mediatica] = 1)
6 VAR Obstetrica = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[obstetrica] = 1)
7 VAR Otras = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[otras] = 1)
8 VAR ValorMaximo = MAXX({Institucional, Laboral, LibertadReproductiva, Mediatica, Obstetrica, Otras}, [Value])
9
10 RETURN
11 SWITCH(
12     TRUE(),
13     ValorMaximo = Institucional, "Institucional",
14     ValorMaximo = Laboral, "Laboral",
15     ValorMaximo = LibertadReproductiva, "Libertad Reproductiva",
16     ValorMaximo = Mediatica, "Mediática",
17     ValorMaximo = Obstetrica, "Obstétrica",
18     ValorMaximo = Otras, "Otras",
19     BLANK()
20 )

```

○ TOTALES

■ TOTAL_VICTIMAS

- Esta medida cuenta el total de las víctimas a partir de la columna “victimID” en la tabla hechos, que funciona como una llave primaria irrepitable.
- `1 TOTAL_VICTIMAS = COUNT(PERFILES_VICTIMAS[victimaID])`

■ TOTAL_VICTIMAS_POR_TIPO_MODALIDAD

- ESTA MEDIDA CAE EN DESUSO DEBIDO A LA IMPLEMENTACION DE LA MEDIDA FALSA: “FALSO_TOTAL_VICTIMAS_POR_TIPO_MODALIDAD”.
- Esta medida cuenta el total de víctimas por modalidad filtrando a partir del tipo de modalidad de la tabla desnormalizada, y gracias a ella se permite una correcta segmentación de las víctimas a la hora de seleccionar un tipo de modalidad en el gráfico de barras del apartado “Modalidades”.

```

1 TOTAL_VICTIMAS_POR_TIPO_MODALIDAD =
2 VAR ModalidadSeleccionada = SELECTEDVALUE(TablaModalidadesDesnormalizada[TipoModalidad])
3 RETURN
4 SWITCH(
5     TRUE(),
6     ModalidadSeleccionada = "Institucional", [INSTITUCIONAL],
7     ModalidadSeleccionada = "Laboral", [LABORAL],
8     ModalidadSeleccionada = "Libertad Reproductiva", [LIBERTAD REPRODUCTIVA],
9     ModalidadSeleccionada = "Mediática", [MEDIATICA],
10    ModalidadSeleccionada = "Obstétrica", [OBSTETRICA],
11    ModalidadSeleccionada = "Otras", [OTRAS],
12    BLANK()
13 )
14

```

■ TOTAL_VICTIMAS_POR_TIPO_VIOLENCIA

- Esta medida cuenta el total de víctimas por violencia filtrando a partir del tipo de violencia de la tabla desnormalizada, y gracias a ella se permite una correcta segmentación de las víctimas a la hora de seleccionar un tipo de modalidad en el gráfico circular del apartado “Violencias”.

```

1 TOTAL_VICTIMAS_POR_TIPO_MODALIDAD =
2 VAR ModalidadSeleccionada = SELECTEDVALUE(TablaModalidadesDesnormalizada[TipoModalidad])
3 RETURN
4 SWITCH(
5     TRUE(),
6     ModalidadSeleccionada = "Institucional", [INSTITUCIONAL],
7     ModalidadSeleccionada = "Laboral", [LABORAL],
8     ModalidadSeleccionada = "Libertad Reproductiva", [LIBERTAD REPRODUCTIVA],
9     ModalidadSeleccionada = "Mediática", [MEDIATICA],
10    ModalidadSeleccionada = "Obstétrica", [OBSTETRICA],
11    ModalidadSeleccionada = "Otras", [OTRAS],
12    BLANK()
13 )

```

■ TOTAL_VICTIMAS_SIN_MODALIDAD

- ESTA MEDIDA CAE EN DESUSO DEBIDO A LA IMPLEMENTACION DE LA MEDIDA FALSA: “FALSO_TOTAL_VICTIMAS_SIN_MODALIDAD”.
- Esta medida calcula el total de víctimas sin modalidad a partir de una variable que luego es sometida a una función condicional ya que la misma es utilizada para una etiqueta de datos en el apartado de “Modalidades” y el objetivo es que muestre “0” al estar influenciada bajo ciertas segmentaciones y no “(En blanco)” o que

simplemente muestre el número total de víctimas sin modalidad en función de dichas segmentaciones.

```
1 TOTAL_VICTIMAS_SIN_MODALIDAD =
2 VAR sinmodalidad = CALCULATE([TOTAL_VICTIMAS], PERFILES_VICTIMAS[tipo_modalidad] = 1)
3 RETURN IF(sinmodalidad = 0, 0, sinmodalidad)
```

■ TOTAL_VICTIMAS_SIN_VIOLENCIA

- Esta medida es exactamente igual que la anterior, con la diferencia de que ésta se utiliza únicamente en el recuadro de texto de “INFORMACION RELEVANTE” en la portada “Inicio”.

```
1 TOTAL_VICTIMAS_SIN_VIOLENCIA =
2 VAR sinviolencia = CALCULATE([TOTAL_VICTIMAS], PERFILES_VICTIMAS[tipo_violencia] = 1)
3 RETURN IF(sinviolencia = 0, 0, sinviolencia)
```

○ VIOLENCIAS

■ PSICOLOGICA

- Esta medida sirve para contar la cantidad total de víctimas por violencia psicológica.

```
1 PSICOLOGICA = CALCULATE([TOTAL_VICTIMAS], TIPOS_VIOLENCIA[psicologica] = 1)
```

■ DOMESTICA

- Esta medida sirve para contar la cantidad total de víctimas por violencia doméstica.

```
1 DOMESTICA = CALCULATE([TOTAL_VICTIMAS], TIPOS_VIOLENCIA[domestica] = 1)
```

■ FISICA

- Esta medida sirve para contar la cantidad total de víctimas por violencia física.

```
1 FISICA = CALCULATE([TOTAL_VICTIMAS], TIPOS_VIOLENCIA[fisica] = 1)
```

■ ECONOMICA-PATRIMONIAL

- Esta medida sirve para contar la cantidad total de víctimas por violencia económica y patrimonial.

```
1 ECONOMICA-PATRIMONIAL = CALCULATE([TOTAL_VICTIMAS], TIPOS_VIOLENCIA[economica_patrimonial] = 1)
```

■ SIMBOLICA

- Esta medida sirve para contar la cantidad total de víctimas por violencia simbólica.

```
1 SIMBOLICA = CALCULATE([TOTAL_VICTIMAS], TIPOS_VIOLENCIA[simbolica] = 1)
```

■ SEXUAL

- Esta medida sirve para contar la cantidad total de víctimas por violencia sexual.

- `1 SEXUAL = CALCULATE([TOTAL_VICTIMAS], TIPOS_VIOLENCIA[sexual] = 1)`

○ MODALIDADES

■ OTRAS

- Esta medida sirve para contar la cantidad total de víctimas por la modalidad “Otras”.

- `1 OTRAS = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[otras] = 1)`

■ INSTITUCIONAL

- Esta medida sirve para contar la cantidad total de víctimas por la modalidad “Institucional”.

- `1 INSTITUCIONAL = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[institucional] = 1)`

■ LABORAL

- Esta medida sirve para contar la cantidad total de víctimas por la modalidad “Laboral”.

- `1 LABORAL = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[laboral] = 1)`

■ LIBERTAD REPRODUCTIVA

- Esta medida sirve para contar la cantidad total de víctimas por la modalidad “Libertad Reproductiva”.

- `1 LIBERTAD REPRODUCTIVA = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[libertad_reproductiva] = 1)`

■ OBSTETRICA

- Esta medida sirve para contar la cantidad total de víctimas por la modalidad “Obstétrica”.

- `1 OBSTETRICA = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[obstetrica] = 1)`

■ MEDIATICA

- Esta medida sirve para contar la cantidad total de víctimas por la modalidad “Mediática”.

- `1 MEDIATICA = CALCULATE([TOTAL_VICTIMAS], TIPOS_MODALIDAD[mediatica] = 1)`

○ OTRAS

○ CANTIDAD_TRIMESTRES

- Esta medida cuenta la cantidad de trimestres filtrando a partir de la fecha máxima del dataset (actualmente 31 de junio del 2023), ya que la tabla calendario abarca hasta el 31 de diciembre del 2023; dicha medida se utiliza en la portada inicio, recuadro de texto “INFORMACION RELEVANTE” y sirve para dividir la cantidad de víctimas por violencia por la cantidad de trimestres hasta la fecha, obteniendo el promedio de víctimas por trimestre. De esta forma en el caso de actualizarse los datos de las

tablas, también se actualizarán los datos en el recaudo de texto del inicio o portada.

```
1 CANTIDAD_TRIMESTRES =
2 VAR fechaMax = MAX(PERFILES_VICTIMAS[fecha_registro])
3 RETURN
4 COUNTROWS(
5     SUMMARIZE(
6         FILTER(CALENDARIO, CALENDARIO[COMPONENTES] <= fechaMax),
7         CALENDARIO[AÑO],
8         CALENDARIO[TRIMESTRE])
9 )
```

○ PROMEDIO_EDAD_VICTIMA

- Esta medida obtiene el promedio de la edad de las víctimas a partir de la tabla de hechos.

```
1 PROMEDIO_EDAD_VICTIMA = AVERAGE(PERFILES_VICTIMAS[edad_victima])
```

○ PROMEDIO_VICTIMAS_POR_AÑO_y_TRIMESTRE

- Esta medida obtiene el promedio de víctimas por trimestre.

```
1 PROMEDIO_VICTIMAS_POR_AÑO_y_TRIMESTRE =
2 VAR anyotrimestre = SUMMARIZE(CALENDARIO, CALENDARIO[COMPONENTES].[Año], CALENDARIO[COMPONENTES].[Trimestre])
3 RETURN AVERAGEX(anyotrimestre, [TOTAL_VICTIMAS])
```

○ PORCENTAJE_VICTIMAS_POR_TIPO_VIOLENCIA

- Esta medida es aplicada en el tooltip “tipos_violencias” (el cual es utilizado en el gráfico de barras de la solapa “VIOLENCIAS” denominado como “COMBINACION DE VIOLENCIAS”) y devuelve el porcentaje de víctimas por tipo de violencia según el número de combinación.

```
1 PORCENTAJE_VICTIMAS_POR_TIPO_VIOLENCIA =
2 VAR ViolenciaSeleccionada = SELECTEDVALUE(TablaViolenciasDesnormalizada[TipoViolencia])
3 VAR TotalVictimas = CALCULATE(COUNT(PERFILES_VICTIMAS[victimaID]), REMOVEFILTERS())
4 RETURN
5     SWITCH(
6         TRUE(),
7         ViolenciaSeleccionada = "Psicológica", DIVIDE([PSICOLOGICA], TotalVictimas),
8         ViolenciaSeleccionada = "Doméstica", DIVIDE([DOMESTICA], TotalVictimas),
9         ViolenciaSeleccionada = "Física", DIVIDE([FISICA], TotalVictimas),
10        ViolenciaSeleccionada = "Económica/Patrimonial", DIVIDE([ECONOMICA-PATRIMONIAL], TotalVictimas),
11        ViolenciaSeleccionada = "Sexual", DIVIDE([SEXUAL], TotalVictimas),
12        ViolenciaSeleccionada = "Simbólica", DIVIDE([SIMBOLICA], TotalVictimas),
13        BLANK()
14    )
```

○ PORCENTAJE_VICTIMAS_POR_TIPO_MODALIDAD

- ESTA MEDIDA CAE EN DESUSO DEBIDO A LA IMPLEMENTACION DE LA MEDIDA FALSA: "FALSO_PORCENTAJE_VICTIMAS_POR_TIPO_MODALIDAD"

- Esta medida es aplicada en el tooltip "tipos_modalidades" (el cual es utilizado en el gráfico de barras de la solapa "MODALIDADES" denominado como "COMBINACION DE MODALIDADES") y devuelve el porcentaje de víctimas por tipo de modalidad según el número de combinación.

```

1 PORCENTAJE_VICTIMAS_POR_TIPO_MODALIDAD =
2 VAR ModalidadSeleccionada = SELECTEDVALUE(TablaModalidadesDesnormalizada[TipoModalidad])
3 VAR TotalVictimas = CALCULATE(COUNT(PERFILES_VICTIMAS[victimaID]), REMOVEFILTERS())
4 RETURN
5 SWITCH(
6     TRUE(),
7     ModalidadSeleccionada = "Institucional", DIVIDE([INSTITUCIONAL], TotalVictimas),
8     ModalidadSeleccionada = "Laboral", DIVIDE([LABORAL], TotalVictimas),
9     ModalidadSeleccionada = "Libertad Reproductiva", DIVIDE([LIBERTAD REPRODUCTIVA], TotalVictimas),
10    ModalidadSeleccionada = "Mediática", DIVIDE([MEDIATICA], TotalVictimas),
11    ModalidadSeleccionada = "Obstétrica", DIVIDE([OBSTETRICA], TotalVictimas),
12    ModalidadSeleccionada = "Otras", DIVIDE([OTRAS], TotalVictimas),
13    BLANK()
14 )
15

```

○ VIOLENCIA_CON_MAS_VICTIMAS

- Esta medida ayuda a determinar el tipo de violencia con más víctimas en el dataset a partir de otras medidas y una variable que utiliza la función "MAXX()" y así tomar el número máximo para luego compararlo y determinar a qué tipo de violencia pertenece.

```

1 VIOLENCIA_CON_MAS_VICTIMAS =
2 VAR Psicologica = [PSICOLOGICA]
3 VAR Domestica = [DOMESTICA]
4 VAR Fisica = [FISICA]
5 VAR EcoPatri = [ECONOMICA-PATRIMONIAL]
6 VAR Simbolica = [SIMBOLICA]
7 VAR Sexual = [SEXUAL]
8 VAR ValorMaximo = MAXX({Psicologica, Domestica, Fisica, EcoPatri, Simbolica, Sexual}, [Value])
9
10 RETURN
11 SWITCH(
12     TRUE(),
13     ValorMaximo = Psicologica, "Psicológica",
14     ValorMaximo = Domestica, "Doméstica",
15     ValorMaximo = Fisica, "Física",
16     ValorMaximo = EcoPatri, "Económica/Patrimonial",
17     ValorMaximo = Simbolica, "Simbólica",
18     ValorMaximo = Sexual, "Sexual",
19     BLANK()
20 )

```

Explicación del Diagrama Obtenido en POWER BI:

- **DESCRIPCION GENERAL**

- El diagrama contiene diez(10) tablas en total, de las cuales nueve(9) son tablas dimensionales que contienen información característica de cada entidad y una(1) es la tabla de hechos que centraliza toda la información y permite conectar los datos y realizar así diversos análisis.

De esta forma se obtiene un esquema en estrella con una tabla de hechos conectada a varias tablas dimensionales.

- La tabla “Medidas” no se tuvo en cuenta para esta descripción.

- **DESCRIPCION DE TABLAS**

- **CALENDARIO**

- Tabla dimensional que proporciona un periodo de tiempo basado en las fechas del dataset.
 - COMPONENTES: Llave foránea (FK) creada a partir de “CALENDARAUOT()” que identifica de manera única cada fecha en la tabla de hechos “PERFILES_VICTIMAS”, columna tipo (DATE).
 - AÑO: Repite el año de la columna “COMPONENTES”, columna tipo (INT).
 - MES: Repite el número del mes de la columna “COMPONENTES”, columna tipo (INT).
 - MES_NOMBRE: Nombre del mes, columna tipo (NVARCHAR).
 - DIA: Repite el número del día de la columna “COMPONENTES” (1 al 31), columna tipo (INT).
 - DIA_ID: Califica un número del 1 al 7 a cada fecha, donde “1” representa el lunes y “7” domingo. columna tipo (INT).
 - DIA_NOMBRE: En función de la columna “DIA_ID” Coloca el nombre del día según corresponda.
 - TRIMESTRE: Indica el número del trimestre (1 al 4) en función de la fecha en la columna “COMPONENTES”

COMPONENTES ▾	AÑO ▾	MES ▾	DIA ▾	DIA_NOMBRE ▾	MES_NOMBRE ▾	DIA_ID ▾	TRIMESTRE ▾
1/1/2020	2020	1	1	Miércoles	Enero	3	1
2/1/2020	2020	1	2	Jueves	Enero	4	1
3/1/2020	2020	1	3	Viernes	Enero	5	1
4/1/2020	2020	1	4	Sábado	Enero	6	1
5/1/2020	2020	1	5	Domingo	Enero	7	1
6/1/2020	2020	1	6	Lunes	Enero	1	1
7/1/2020	2020	1	7	Martes	Enero	2	1
8/1/2020	2020	1	8	Miércoles	Enero	3	1
9/1/2020	2020	1	9	Jueves	Enero	4	1
10/1/2020	2020	1	10	Viernes	Enero	5	1
11/1/2020	2020	1	11	Sábado	Enero	6	1
12/1/2020	2020	1	12	Domingo	Enero	7	1
13/1/2020	2020	1	13	Lunes	Enero	1	1
14/1/2020	2020	1	14	Martes	Enero	2	1

PAISES

- Tabla dimensional que contiene de manera única los distintos países existentes en el dataset; contiene cuarenta y dos(42) posibilidades.
 - paisID: Llave primaria (PK) para identificar de manera única cada país, columna tipo (INT).
 - nombre_pais: Nombre del país, columna tipo (NVARCHAR).

paisID ▾	nombre_pais ▾
1	Alemania
2	Arabia
3	Argentina
4	Armenia
5	Australia
6	Bélgica
7	Bolivia
8	Brasil
9	Canadá
10	Chile
11	Colombia
12	Costa Rica
13	Cuba
14	Desconocido
15	Ecuador
16	El Salvador

PROVINCIAS

- Tabla dimensional que contiene de manera única las distintas provincias existentes en el dataset junto con la población de cada una registrada hasta entonces; contiene veinticinco(25) posibilidades.
 - provincialID: Llave primaria (PK) para identificar de manera única cada provincia, columna tipo (INT).
 - nombre_provincia: Nombre de la provincia, columna tipo (NVARCHAR).

provincialID	nombre_provincia
1	Ciudad Autónoma de Buenos Aires
2	San Luis
3	San Juan
4	Buenos Aires
5	Misiones
6	Córdoba
7	Salta
8	Santiago Del Estero
9	Tucumán
10	Jujuy
11	Desconocido
12	La Pampa
13	La Rioja
14	Catamarca

- **TablaCoordenadas**

- Tabla dimensional que contiene coordenadas específicas de cada provincia del dataset
 - Provincia: Llave foránea (FK) que contiene los nombres de las provincias y que conecta con la tabla "PROVINCIAS" para identificar la provincia, columna tipo (INT).
 - Latitud: Latitud de la provincia, columna tipo (FLOAT).
 - Longitud: Longitud de la provincia, columna tipo (FLOAT).
 - Provincia_Pais: Contiene el nombre de las provincias concatenado a la palabra "Argentina" para evitar errores en el gráfico de mapa, columna tipo (NVARCHAR).

Provincia	Latitud	Longitud	Provincia_Pais
Ciudad Autónoma de Buenos Aires	-346037	-583816	Ciudad Autónoma de Buenos Aires, Argentina
Buenos Aires	-349214	-579544	Buenos Aires, Argentina
San Luis	-333018	-663378	San Luis, Argentina
Catamarca	-284696	-657795	Catamarca, Argentina
San Juan	-315375	-685364	San Juan, Argentina
Chaco	-274512	-589867	Chaco, Argentina
Chubut	-433007	-651022	Chubut, Argentina
Misiones	-273668	-558961	Misiones, Argentina
Córdoba	-314201	-641888	Córdoba, Argentina
Corrientes	-274692	-588306	Corrientes, Argentina
Salta	-247821	-654232	Salta, Argentina
Entre Ríos	-317333	-605333	Entre Ríos, Argentina
Santiago Del Estero	-277834	-642642	Santiago Del Estero, Argentina
Tucumán	-268083	-652176	Tucumán, Argentina
Jujuy	-241858	-652995	Jujuy, Argentina
La Pampa	-366206	-6429	La Pampa, Argentina

TIPOS_VIOLENCIA

- Tabla dimensional que posee los diferentes tipos de violencia existentes en el dataset; contiene sesenta y tres (63) combinaciones.
 - tipovienciaID: Llave primaria (PK) para identificar de manera única cada tipo de violencia, columna tipo (INT).
 - fisica: Indicador binario de violencia física, columna tipo (INT).
 - psicologica: Indicador binario de violencia psicológica, columna tipo (INT).
 - sexual: Indicador binario de violencia sexual, columna tipo (INT).
 - economica_patrimonial: Indicador binario de violencia económica y patrimonial, columna tipo (INT).
 - simbolica: Indicador binario de violencia simbólica, columna tipo (INT).
 - domestica: Indicador binario de violencia doméstica, columna tipo (INT).

tipovienciaID	fisica	psicologica	sexual	economica_patrimonial	simbolica	domestica
1	0	0	0	0	0	0
2	1	1	0	1	0	1
3	1	0	1	1	1	1
4	1	0	0	1	1	1
5	1	0	1	1	0	0
6	0	0	1	1	1	0
7	1	1	1	0	1	1
8	0	0	1	1	0	1
9	1	0	0	0	0	1
10	0	0	0	1	0	1
11	0	1	1	1	1	1
12	1	0	1	0	1	0
13	1	0	1	0	0	1
14	1	0	0	1	0	0
15	1	1	1	1	0	1

○ TablaViolenciasDesnormalizada

- Tabla dimensional que contiene detalles desnormalizados o redundantes sobre la tabla “TIPOS_VIOLENCIA” que confirma si las diferentes combinaciones poseen o no los diferentes tipos de violencia.
 - tipovienciaID: Llave foránea (FK) para identificar de manera única cada tipo de violencia, columna tipo (INT).
 - TipoViolencia: Descripción textual del tipo de violencia, columna tipo (NVARCHAR).
 - Valor: Indica si una combinación posee o no un tipo de violencia, columna tipo (INT)

tipovienciaID	TipoViolencia	Valor
1	Física	0
1	Sexual	0
1	Económica/Patrim	0
1	Psicológica	0
1	Simbólica	0
1	Doméstica	0
2	Económica/Patrim	1
2	Psicológica	1
2	Física	1
2	Sexual	0
2	Simbólica	0
2	Doméstica	1
3	Psicológica	0
3	Física	1
3	Sexual	1
3	Económica/Patrim	1
3	Simbólica	1
3	Doméstica	1

TIPOS_MODALIDAD

- Tabla dimensional que posee los diferentes tipos de modalidades de violencia existentes en el dataset; contiene veintidós(22) combinaciones.
 - tipomodalidadID: Llave primaria (PK) para identificar de manera única cada tipo de modalidad de violencia; columna tipo (INT).
 - institucional: Indicador binario de violencia institucional, columna tipo (INT).
 - laboral: Indicador binario de violencia laboral, tipo (INT).
 - libertad_reproductiva: Indicador binario de violencia contra la libertad reproductiva, columna tipo (INT).
 - obstetrica: Indicador binario de violencia obstétrica, tipo (INT).
 - mediatica: Indicador binario de violencia mediática, columna tipo (INT).
 - otras: Indicador binario de otras modalidades de violencia, columna tipo (INT).

tipomodalidadID	institucional	laboral	libertad_reproductiva	obstetrica	mediatica	otras
1	0	0	0	0	0	0
2	1	0	1	1	0	0
3	1	0	0	0	0	1
4	0	0	0	1	0	1
5	1	0	0	1	0	0
6	1	1	0	0	0	0
7	1	1	1	0	0	0
8	0	0	0	0	0	1
9	1	1	0	1	0	0
10	1	0	0	0	1	0
11	0	1	1	0	0	0
12	1	0	1	0	0	0
13	0	1	0	0	0	0
14	1	0	0	0	0	0

○ TablaModalidadesDesnormalizada

- Tabla dimensional que contiene detalles desnormalizados o redundantes sobre la tabla “TIPOS_MODALIDAD” que confirma si las diferentes combinaciones poseen o no los diferentes tipos de violencia.
 - tipovienciaID: Llave foránea (FK) para identificar de manera única cada tipo de modalidad, columna tipo (INT).
 - TipoModalidad: Descripción textual del tipo de modalidad, columna tipo (NVARCHAR).
 - Valor: Indica si una combinación posee o no un tipo de modalidad, columna tipo (INT).

tipomodalidadID	TipoModalidad	Valor
1	Libertad Reproductiva	0
1	Mediática	0
1	Institucional	0
1	Laboral	0
1	Obstétrica	0
1	Otras	0
2	Laboral	0
2	Mediática	0
2	Otras	0
2	Institucional	1
2	Libertad Reproductiva	1
2	Obstétrica	1
3	Obstétrica	0
3	Laboral	0
3	Libertad Reproductiva	0
3	Mediática	0
3	Institucional	1
3	Otras	1

○ PERFILES_AGRESOR

- Tabla dimensional que contiene los diferentes tipos de perfiles de agresores existentes del dataset; contiene cuarenta y tres(43) combinaciones.
 - tipoagresorID: Llave primaria (PK) para identificar de manera única cada perfil de agresor, columna tipo (INT).
 - genero_agresor: Género del agresor, columnas tipo (NVARCHAR).
 - vinculo_agresor: Vínculo del agresor con la víctima, columna tipo (NVARCHAR).

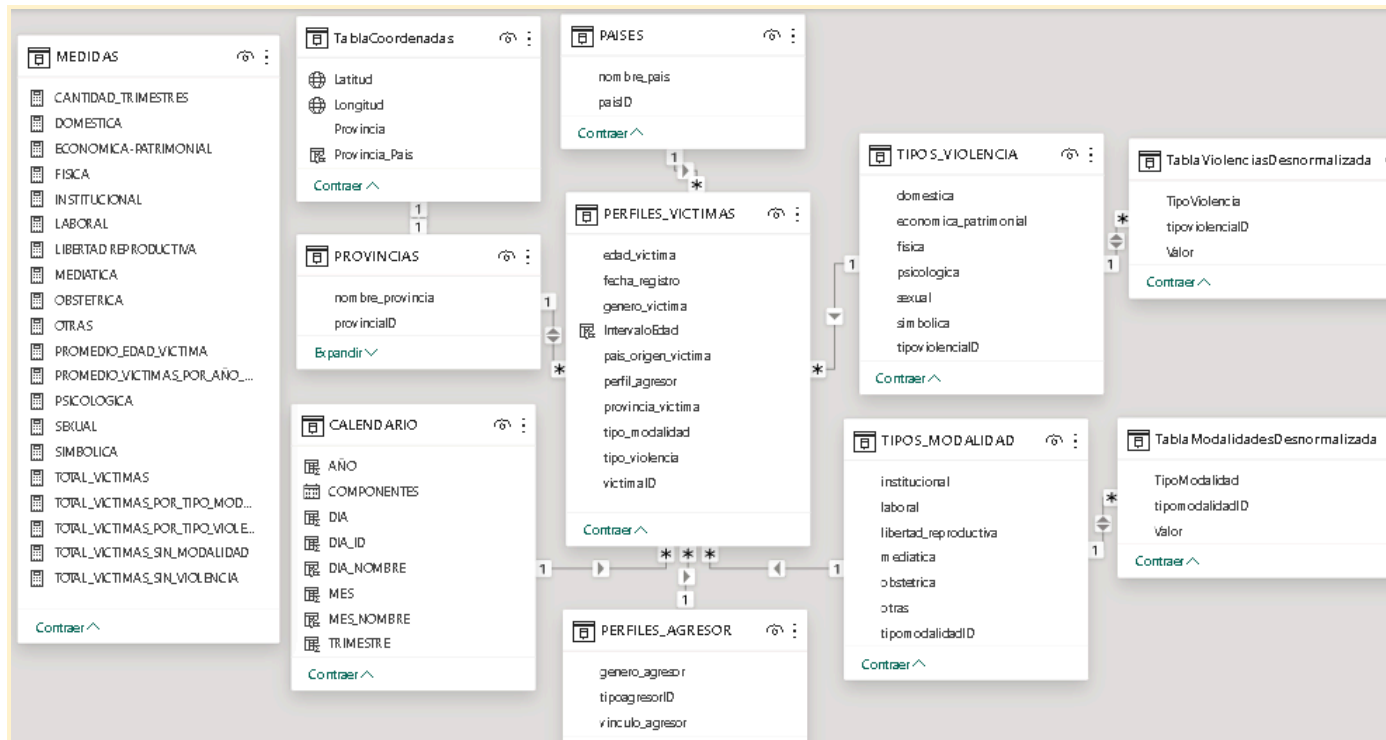
tipoagresorID	genero_agresor	vinculo_agresor
1	Varon Trans	Pareja
2	Varon	Superior jerarquico
3	Varon	Otro familiar
4	Varon	Ex pareja
5	Transgenero	Pareja
6	Mujer Trans	Ex pareja
7	Otro	Otro familiar
8	Otro	Superior jerarquico
9	Mujer	Madre o tutor
10	Desconocido	Otro
11	Otro	Ex pareja
12	Desconocido	Madre o tutor
13	Desconocido	Ex pareja
14	Desconocido	Padre o tutor

PERFILES_VICTIMAS | Tabla de Hechos

- Tabla de hechos que conecta las características de las diferentes entidades en el esquema de datos que están relacionadas con las víctimas y los incidentes de violencia de género; contiene ochentinueve mil quinientos cuarenta(89540) registros/filas.
 - victimaID: Llave primaria (PK) para identificar de manera única cada víctima, columna tipo (INT).
 - fecha_registro: Fecha en la que se registró el caso, columna tipo (DATE).
 - pais_origen_victima: Llave foránea (FK) de la tabla "PAISES" para identificar el país de origen de la víctima, columna tipo (INT).
 - provincia_victima: Llave foránea (FK) de la tabla "PROVINCIAS" para identificar la provincia de la víctima, columna tipo (INT).
 - genero_victima: Género de la víctima, columna tipo (NVARCHAR).
 - edad_victima: Edad de la víctima, columna tipo (INT).
 - IntervaloEdad: Cataloga a las víctimas por intervalo de edad según su edad, columna tipo (NVARCHAR).
 - tipo_violencia: Llave foránea (FK) de la tabla "TIPOS_VIOLENCIA" para identificar el tipo de violencia sufrida por la víctima, columna tipo (INT).
 - tipo_modalidad: Llave foránea (FK) de la tabla "TIPOS_MODALIDAD" para identificar el tipo de modalidad sufrida por la víctima, columna tipo (INT).
 - perfil_agresor: Llave foránea (FK) de la tabla "PERFILES_AGRESOR" para identificar el tipo de perfil del agresor, columna tipo (INT).

victimaID	fecha_registro	pais_origen_victima	provincia_victima	genero_victima	edad_victima	tipo_violencia	tipo_modalidad	perfil_agresor	IntervaloEdad
1	1/1/2023	14	4	Mujer	35	34	1	26	31-40
2	1/1/2023	3	4	Mujer	19	25	1	28	11-20
3	1/1/2023	3	4	Mujer	46	26	1	26	41-50
4	1/1/2023	14	24	Mujer	28	33	1	26	21-30
5	1/1/2023	3	20	Mujer	34	56	1	4	31-40
6	1/1/2023	34	4	Mujer	38	2	1	4	31-40
7	1/1/2023	3	4	Mujer	57	2	1	26	51-60
8	1/1/2023	3	9	Mujer	35	26	1	26	31-40
9	1/1/2023	3	16	Mujer	39	42	1	4	31-40
10	1/1/2023	3	20	Mujer	24	15	1	26	21-30
11	1/1/2023	14	4	Mujer	35	34	1	4	31-40
12	1/1/2023	3	5	Mujer	22	2	1	26	21-30
13	1/1/2023	14	16	Mujer	52	34	1	26	51-60
14	1/1/2023	3	16	Mujer	65	2	1	4	61-70
15	1/1/2023	14	4	Mujer	26	2	1	26	21-30
16	1/1/2023	14	16	Mujer	41	2	1	26	41-50

● **DIAGRAMA RESULTANTE EN POWER BI:**



Herramientas Utilizadas:

- Pandas v2.2.2
 - Utilizada para la lectura de registros nulos.
- SQL Server v16.0.1000.6
 - Utilizada para la lectura, limpieza y transformación de datos.
- Power BI v2.124.2028.0 32-bit (diciembre de 2023)
 - Utilizada para la realización del dashboard y el análisis de datos.
- Power Point v2408 (compilación 16.0.17928.20114)
 - Utilizada para la creación de marcadores.

ANEXO

[CARPETA DRIVE CON DICHOS ARCHIVOS](#)