



**TRABAJO DE GRADO**  
**Opción Seminario-Diplomado.**

Violencia Doméstica en Bogotá y Medellín en el año 2021 mediante la Metodología del

Análisis de Datos

Corporación Universitaria Remington

Facultad Ingenierías

Ingeniería de Sistemas

Alejandro López Vélez - Juan Camilo Jiménez

Ivonne Castaño Osorio

Jhon Edison Amortegui Granada

Opción de Trabajo de grado Seminario de Grado

2024.

Noveno Semestre

Ingeniería de Sistemas

[Alejandro.lopez8695@miremington.edu.co](mailto:Alejandro.lopez8695@miremington.edu.co)

[juan.jimenez.1114@miremington.edu.co](mailto:juan.jimenez.1114@miremington.edu.co)

## **Dedicatoria**

Este trabajo de grado está dedicado a unas personas muy especiales sin las cuales esto no hubiera sido posible.

A mi familia, porque sin su amor, motivación y apoyo no hubiera tenido las fuerzas para culminar toda la Carrera y finalizar con éxito mis objetivos. A mi mamá y a mi papa que siempre creyeron en mí y me brindaron todo su acompañamiento.

## Tabla de contenidos

|   |    |
|---|----|
| Resumen.....  | 4  |
| Palabras clave: .....   | 4  |
| Análisis temporal-factores de riesgo- prevención-violencia intrafamiliar-patrones.....        | 4  |
| Pregunta orientadora de la búsqueda .....   | 5  |
| Metodología de búsqueda de la información .....   | 6  |
| Comprender su negocio: .....  | 6  |
| Métodos de análisis:.....   | 6  |
| Requisitos de datos: .....  | 6  |
| Recopilación de datos: .....  | 6  |
| Comprensión de datos:.....  | 7  |
| Preparación de datos: .....   | 7  |
| Modelado: .....   | 7  |
| Evaluación: .....   | 8  |
| Implementación: .....   | 8  |
| Sustentación teórica de la pregunta.....  | 9  |
| Figuras y tablas .....  | 10 |
| Gráfico 1. Bigotes. ....  | 10 |
| Gráfico2. Histograma. ....  | 11 |
| Gráfico 3. Dispersión. ....   | 12 |
| Tabla 1. Estadísticas Descriptivas de Variables Clave en Violencia Intrafamiliar (2021). .... | 13 |
| Tabla 2. Correlaciones entre Variables de Violencia Intrafamiliar (2021).....                 | 14 |
| Conclusiones. ....  | 15 |
| Lista de referencias .....  | 16 |

## **Resumen**

Este estudio examina los incidentes de vandalismo en Colombia en 2021, centrándose en el análisis de datos como el plazo de los incidentes, el individuo, la variedad de las víctimas, la perduración y la frecuencia de los ataques por mes. El objetivo es identificar patrones de tiempo, factores de riesgo y comportamientos asociados a estos eventos. Este ejercicio tiene como objetivo proporcionar información importante para comprender mejor las condiciones básicas, las consecuencias de las familias afectadas y desarrollar estrategias efectivas para prevenir y reducir la violencia doméstica en el contexto de Colombia.

### **Palabras clave:**

Análisis temporal-factores de riesgo- prevención-violencia intrafamiliar-patrones.

### **Pregunta orientadora de la búsqueda**

Utilizando técnicas de análisis de datos, el estudio de violencia intrafamiliar en Bogotá y Medellín durante 2021 busca identificar las áreas con las tasas más altas de incidentes. La investigación examina cómo varían estos incidentes a lo largo del año según la edad y el género de las víctimas, así como las leyes locales y las medidas preventivas afectaron a las víctimas. Los datos muestran una gran variación en la cantidad y la fecha de los incidentes de violencia con armas, según el grupo etario. Este enfoque ayuda en la identificación de patrones demográficos y temporales significativos, lo que ayuda a la creación de políticas e intervenciones más efectivas en ambas ciudades.

¿En dónde se encuentran reflejados las mayores tasas de violencia intrafamiliar en Bogotá y Medellín en el año 2021, utilizando metodología del análisis de datos?

## **Metodología de búsqueda de la información**

### **Comprender su negocio:**

Este proceso comienza tomándose el tiempo para recopilar información para comprender de qué se trata su negocio y cómo funciona. Entender el problema de la empresa o lo que busca es clave para crear buenas preguntas.

### **Métodos de análisis:**

Luego de identificar el problema o necesidad de la empresa, seleccionar un método de análisis en función de las necesidades del cliente y de los responsables de interpretar las técnicas más efectivas capítulo a los problemas se les asignan diagnósticos basados en el comportamiento observado y se confirman mediante modelos descriptivos o predictivos.

### **Requisitos de datos:**

Importante analizar nuestros datos y evitar investigaciones centrándose en la investigación y el análisis basados en datos reales.

### **Recopilación de datos:**

La primera recolección de datos es para verificar el cumplimiento; si tiene datos que contienen la información correcta, puede tener más ideas sobre qué analizar según la pregunta planteada. En este punto es preciso definir qué se necesita en el análisis.

Para este estudio, se utilizó la base de datos "Domestic violence in Colombia" disponible en Kaggle, la cual da información detallada de los incidentes de violencia doméstica en Colombia en el año 2021, esta base de datos fue esencial para identificar los patrones. (BIO, 2021)

**Comprensión de datos:**

Esta sección utiliza técnicas de análisis gráfico para ayudar a comprender cómo se relacionan las diferentes variables entre sí. Este paso utiliza soporte de histograma para organizar los datos para un mejor análisis y lectura. Este paso es muy importante ya que requiere procesamiento y distribución de los datos. Minimiza el proceso de búsqueda y reduce significativamente el tiempo de prueba. Compruebe si hay duplicaciones o errores.

**Preparación de datos:**

La fase de preparación de datos cubre todos los pasos para construir el conjunto de datos final a partir del conjunto de datos original [los datos deben proporcionarse en el resumen]. Las operaciones de preparación de datos se pueden realizar varias veces sin secuencia. Las responsabilidades incluyen seleccionar tablas, archivos y propiedades, así como editar y limpiar materiales de herramientas de presentación (Chapman, 2000). Durante esta fase, el científico de datos es responsable de limpiar los datos proporcionados, editarlos y prepararlos para su presentación en esta fase, crear nuevas variables o modificar las variables existentes para respaldar el análisis de los datos dados.

Se transfieren los datos obtenidos en formato CSV a Excel y use Power Query para verificarlos, determinar si hay errores en las columnas, corregir estos errores y analizar las columnas donde los datos son inapropiados. Se cambió el formato de fecha, se corrigieron celdas con formato incorrecto, se cambiaron las celdas deseadas a números en el texto. Esto es para limpiar completamente la información que se está analizando.

**Modelado:**

En esta sección, se construye y se prueban diferentes modelos para analizar la información proporcionada y se elige el tipo de modelo que informa, ya sea descriptivo, analítico o compositivo, todo con el objetivo de demostrar veracidad y brillo; Los datos se analizan mediante técnicas como la regresión lineal, que predicen el comportamiento de variables (opiniones) en función de variables específicas (Dagnino, 2014). Además se utilizan técnicas de análisis descriptivo para estudiar el conjunto de datos. Este método tiene como objetivo hacer inferencias o predicciones sobre personas a partir de una muestra (Saldívar, 2008).

### **Evaluación:**

La evaluación se lleva a cabo durante la creación del prototipo y antes de la implementación, lo que ayuda a probar la calidad del diseño del prototipo y puede comprobar si es relevante para la aplicación original. Además, intenta explicar los resultados obtenidos observando patrones y relaciones entre todos los tipos.

### **Implementación:**

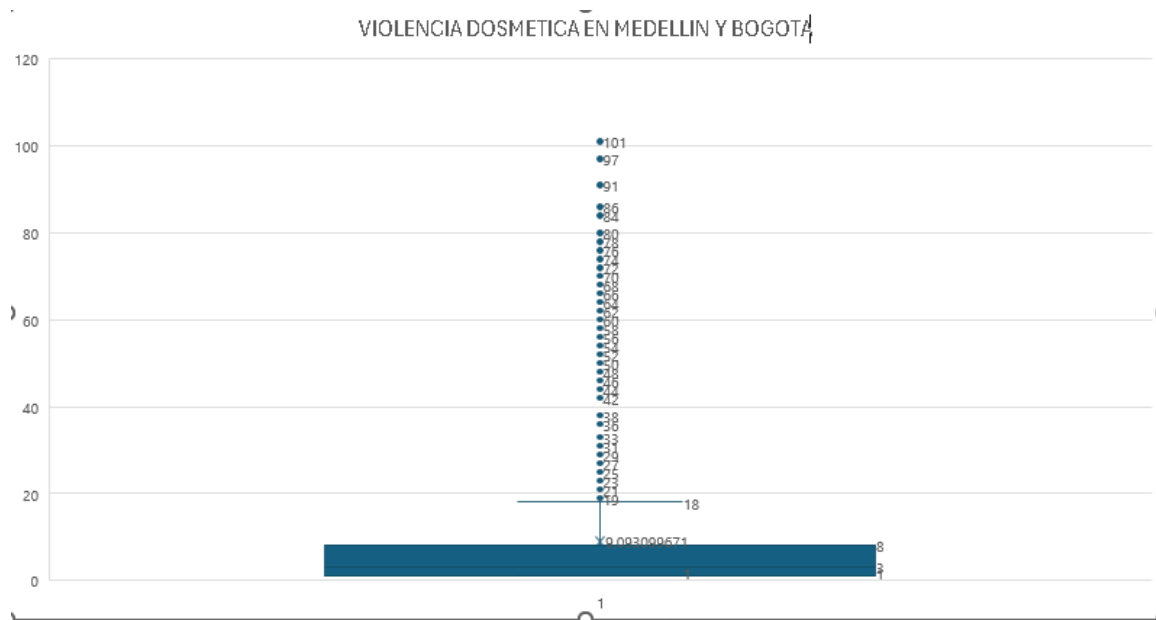
En esta fase, una vez establecido y validado el diseño, la información obtenida se traduce en acciones en el proceso de negocio, por ejemplo, el investigador recomienda acciones basadas en el modelo y la gama de datos que genera (Cortina, 2015). Comentario Esta parte del método CRISP-DM no se utilizó durante el análisis de datos.

## Sustentación teórica de la pregunta

Este análisis se centra en información confiable para comprender las tendencias que se muestran según el informe realizado anteriormente teniendo en cuenta que se manifiestan mucho más una frecuencia de violencia intrafamiliar baja, se denota también que hay picos de violencia intrafamiliar alta y la variedad no es predecible, se precisa investigar las tendencias de mayor gravedad de violencia que son en momentos específicos o causadas por algún factor externo. se tiene ya planteado que podría ser por causales socioeconómicas de estrés, ansiedad, depresión, drogas y conflicto armado dentro de las mismas comunidades y en los estratos más bajos que es donde mayormente se presentaría este tipo de violencia.

También se tienen en cuenta la edad y el sexo de las víctimas, ya que pueden influir en la frecuencia y distribución de la violencia doméstica. Considerando las condiciones socioeconómicas y culturales de Bogotá y Medellín, es importante examinar si la incidencia de la violencia familiar varía según estas condiciones (Salud Mental, 2024).

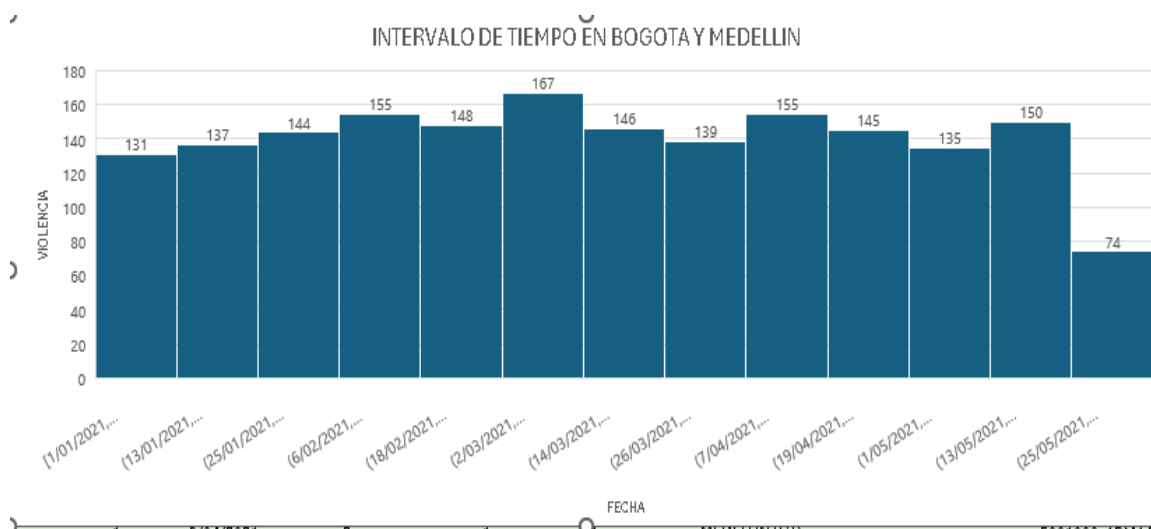
## Figuras y tablas



**Gráfico 1. Bigotes.**

(López & Jiménez, 2024)

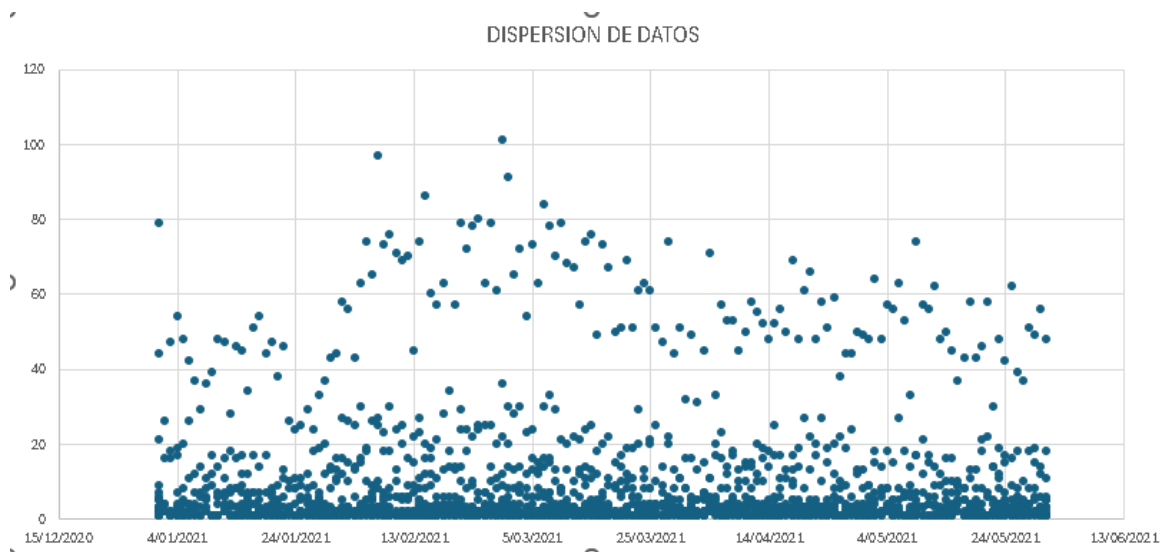
Caja y bigotes, en el gráfico se observa que los valores atípicos se extienden hasta los 101 incidentes con una media general alrededor de 18 incidentes por periodo, la barra central sugiere que los valores concentrados son de bajo valor con algunas alzas de violencia con alta variable significativa de valores atípicos en altos grados.



### Gráfico2. Histograma.

(López & Jiménez, 2024)

Según el gráfico de barras en un intervalo de tiempo del 14 de marzo del 2021 con 167 casos de violencia doméstica siendo este el valor más relevante y el valor más bajo siendo 74 incidentes en la fecha del 25 de mayo del 2021.



### **Grafico 3. Dispersión.**

(López & Jiménez, 2024)

Gráfico de datos dispersos la gráfica muestra las fluctuaciones de incidentes que representan picos por debajo de los 20 incidentes, pero hay picos que superan los 60 incidentes a lo largo del tiempo, la gran mayoría de incidentes son de baja intensidad.

**Tabla 1. Estadísticas Descriptivas de Variables Clave en Violencia Intrafamiliar (2021).**

| CODIGO DANE    |              | ARMAS MEDIOS              |           | FECHA HECHO    |              | GENERO         |           | GRUPO ETARIO   |           | CANTIDAD       |           |
|----------------|--------------|---------------------------|-----------|----------------|--------------|----------------|-----------|----------------|-----------|----------------|-----------|
| Media          | 37751007,45  | Media                     | 3,793888  | Media          | 42841,69632  | Media          | 1,7705698 | Media          | 1,172442  | Media          | 1,7065541 |
| Error típico   | 39884,5689   | Error típico              | 0,0025583 | Error típico   | 1,58474296   | Error típico   | 0,000612  | Error típico   | 0,000737  | Error típico   | 0,0048373 |
| Mediana        | 25754000     | Mediana                   | 3         | Mediana        | 42998        | Mediana        | 2         | Mediana        | 1         | Mediana        | 1         |
| Moda           | 11001000     | Moda                      | 3         | Moda           | 43831        | Moda           | 2         | Moda           | 1         | Moda           | 1         |
| Desviación e   | 27495919,62  | Desviación estándar       | 1,763653  | Desviación e   | 1092,501843  | Desviación e   | 0,4219255 | Desviación e   | 0,5072144 | Desviación e   | 3,3347848 |
| Varianza de la | 7,56026E+14  | Varianza de la muestra    | 3,1104719 | Varianza de la | 1193560,278  | Varianza de la | 0,1780212 | Varianza de la | 0,2572665 | Varianza de la | 11,12079  |
| Curtosis       | -1,480002466 | Curtosis                  | 1,1528018 | Curtosis       | -0,632595461 | Curtosis       | -0,329151 | Curtosis       | 7,3529017 | Curtosis       | 215,88632 |
| Coficiente c   | 0,315724671  | Coficiente de asimetría   | 1,1433659 | Coficiente c   | -0,590202277 | Coficiente c   | -1,26244  | Coficiente c   | 2,922481  | Coficiente c   | 12,601571 |
| Rango          | 94772000     | Rango                     | 9         | Rango          | 4168         | Rango          | 2         | Rango          | 3         | Rango          | 129       |
| Mínimo         | 5001000      | Mínimo                    | 1         | Mínimo         | 40179        | Mínimo         | 1         | Mínimo         | 1         | Mínimo         | 1         |
| Máximo         | 99773000     | Máximo                    | 10        | Máximo         | 44347        | Máximo         | 3         | Máximo         | 4         | Máximo         | 130       |
| Suma           | 1,79414E+13  | Suma                      | 1803087   | Suma           | 20360730384  | Suma           | 841481    | Suma           | 555333    | Suma           | 811062    |
| Cuenta         | 475255       | Cuenta                    | 475261    | Cuenta         | 475255       | Cuenta         | 475260    | Cuenta         | 473655    | Cuenta         | 475263    |
| Mayor(1)       | 99773000     | Mayor(1)                  | 10        | Mayor(1)       | 44347        | Mayor(1)       | 3         | Mayor(1)       | 4         | Mayor(1)       | 130       |
| Menor(1)       | 5001000      | Menor(1)                  | 1         | Menor(1)       | 40179        | Menor(1)       | 1         | Menor(1)       | 1         | Menor(1)       | 1         |
| Nivel de conf  | 86553,37913  | Nivel de confianza(97,0%) | 0,0055517 | Nivel de conf  | 3,439045777  | Nivel de conf  | 0,0013282 | Nivel de conf  | 0,0015993 | Nivel de conf  | 0,0104974 |

(López & Jiménez, 2024)

En la tabla estadísticas descriptivas de variables clave de violencia intrafamiliar en Colombia en 2021, como Código DANE, Armas Medios, Fecha Hecho, Género, Grupo Etario y Cantidad. Se destacan medidas como la media, la mediana, y la desviación estándar. Por ejemplo, la media de Armas Medios es 3.79.

**Tabla 2. Correlaciones entre Variables de Violencia Intrafamiliar (2021).**

| A            | B            | C            | D           | E           | F            | G          |
|--------------|--------------|--------------|-------------|-------------|--------------|------------|
|              | CODIGO DANE  | ARMAS MEDIOS | FECHA HECHO | GENERO      | GRUPO ETARIO | CANTIDAD   |
| CODIGO DANE  | 7,56024E+14  |              |             |             |              |            |
| ARMAS MEDIOS | -5594747,527 | 3,110465378  |             |             |              |            |
| FECHA HECHO  | -4679227466  | 381,8912296  | 1193557,77  |             |              |            |
| GENERO       | 380766,7993  | -0,005023354 | -4,73049242 | 0,17802079  |              |            |
| GRUPO ETARIO | -677671,8511 | 0,056278368  | 26,5522004  | -0,03425295 | 0,25726591   |            |
| CANTIDAD     | -8035460,825 | 0,35611213   | -12,060527  | 0,0606372   | -0,06233423  | 11,1207665 |

(López & Jiménez, 2024)

En la tabla presenta conexión entre las variables del estudio. Por ejemplo, hay una correlación positiva entre Armas Medios y Fecha Hecho (3.11) y una negativa entre Fecha Hecho y Género (-4.73). Estas relaciones ayudan a entender mejor la dinámica de la violencia intrafamiliar.

## **Conclusiones.**

La violencia intrafamiliar ha alcanzado cifras alarmantes en los últimos años, especialmente en 2021. Solo en las ciudades de Medellín y Bogotá, se registraron más de 300,000 casos. Esta tendencia al alza es evidente durante fechas especiales, exacerbando el problema. Bogotá, con una población de más de 7 millones de habitantes, enfrenta una situación preocupante donde el 52% son mujeres, quienes predominantemente son víctimas de violencia intrafamiliar.

Durante 2021 y 2022, la administración ha mostrado una negligencia palpable. La ineficiencia en los canales de atención, el recorte de programas sociales y la falta de liderazgo de los alcaldes han contribuido a un aumento del 200% en incidentes en comparación con años anteriores. A pesar de los anuncios de la alcaldesa de Bogotá sobre una inversión de más de 12,000 millones de pesos para la prevención de estos delitos, los fondos terminan en manos de alcaldes locales, quienes no han logrado contener la crisis.

En Medellín, también se ha observado un incremento del 66% en casos de violencia intrafamiliar. Es evidente que ninguna localidad, departamento o municipio ha manejado adecuadamente esta situación, demostrando una negligencia generalizada por parte de los organismos de prevención del delito.

Los datos recopilados muestran un alarmante aumento en el índice de violencia en el país. Cada hora mueren tres mujeres, totalizando 128 diariamente y 47,000 a lo largo del año como víctimas de violencia intrafamiliar. Es crucial que como sociedad, denunciemos cualquier forma de abuso y trabajemos juntos para mitigar este grave problema.

## Lista de referencias

BIO. (2021). *kaggle*. Obtenido de Domestic violence in Colombia:

<https://www.kaggle.com/datasets/oscardavidperilla/domestic-violence-in-colombia>

Chapman, P. (2000). *CRISP-DM 1.0*.

Cortina, G. (octubre de 2015). *Aplicacion de la Metodologia CRISP-DM a un Proyecto de .*

Obtenido de Universidad Carlos III de Madrid:

[https://www.academia.edu/download/49071533/PFC\\_Victor\\_Galan\\_Cortina\\_DM.pdf](https://www.academia.edu/download/49071533/PFC_Victor_Galan_Cortina_DM.pdf)

Dupret, M. A. (2014). *https://elibro.net/*. Obtenido de La violencia intrafamiliar contra niños, niñas y adolescentes ¿cómo intervenir y cómo prevenir?:

[https://elibro.net/es/ereader/remington/79855?as\\_all=violencia\\_\\_intrafamiliar&as\\_all\\_op=unaccent\\_\\_icontains&prev=as](https://elibro.net/es/ereader/remington/79855?as_all=violencia__intrafamiliar&as_all_op=unaccent__icontains&prev=as)

López, A., & Jiménez, J. C. (2024).

MENTAL, S. (27 de mayo de 2024). *asi vamos en salud*. Obtenido de

[https://www.asivamosensalud.org/indicadores/salud-mental/tasa-de-violencia-](https://www.asivamosensalud.org/indicadores/salud-mental/tasa-de-violencia-intrafamiliar-georeferenciado#:~:text=En%20el%202019%20se%20report%C3%B3,violencia%20intrafamiliar%20ocurridos%20en%20Colombia.)

[intrafamiliar-](https://www.asivamosensalud.org/indicadores/salud-mental/tasa-de-violencia-intrafamiliar-georeferenciado#:~:text=En%20el%202019%20se%20report%C3%B3,violencia%20intrafamiliar%20ocurridos%20en%20Colombia.)

[georeferenciado#:~:text=En%20el%202019%20se%20report%C3%B3,violencia%20intrafamiliar%20ocurridos%20en%20Colombia.](https://www.asivamosensalud.org/indicadores/salud-mental/tasa-de-violencia-intrafamiliar-georeferenciado#:~:text=En%20el%202019%20se%20report%C3%B3,violencia%20intrafamiliar%20ocurridos%20en%20Colombia.)

[afamiliar%20ocurridos%20en%20Colombia.](https://www.asivamosensalud.org/indicadores/salud-mental/tasa-de-violencia-intrafamiliar-georeferenciado#:~:text=En%20el%202019%20se%20report%C3%B3,violencia%20intrafamiliar%20ocurridos%20en%20Colombia.)

Saldívar, M. (2008). *Instituto Politécnico Nacional*. Obtenido de Estadística descriptiva:

<https://elibro.net/es/ereader/remington/74722?page=1>

SALUD MENTAL. (27 de mayo de 2024). *asi vamos en salud*. Obtenido de asi vamos en salud:

[https://www.asivamosensalud.org/indicadores/salud-mental/tasa-de-violencia-](https://www.asivamosensalud.org/indicadores/salud-mental/tasa-de-violencia-intrafamiliar-georeferenciado#:~:text=En%20el%202019%20se%20report%C3%B3,violencia%20intrafamiliar%20ocurridos%20en%20Colombia.)

[intrafamiliar-](https://www.asivamosensalud.org/indicadores/salud-mental/tasa-de-violencia-intrafamiliar-georeferenciado#:~:text=En%20el%202019%20se%20report%C3%B3,violencia%20intrafamiliar%20ocurridos%20en%20Colombia.)

