Tablero de control para la identificación de patrones y tendencias de la violencia doméstica en Colombia

Adriana Patricia Fontalvo Pérez
Álvaro Andrés Basto Gordillo
Jerson Gustavo Gómez Pérez
Karina Andrea Castro Molina
William Iván Cardón Sotaquirá

BIGDATA, Fundación Universitaria Compensar
Proyecto fin de grado Especialización de Big Data

Omar Alexander León García

29 de mayo de 2023

Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional.

Tablero de control para la identificación de patrones y tendencias de la violencia doméstica en Colombia

Adriana Patricia Fontalvo Pérez
Álvaro Andrés Basto Gordillo
Jerson Gustavo Gómez Pérez
Karina Andrea Castro Molina
William Iván Cardón Sotaquirá

Trabajo de grado presentado como requisito parcial para optar al título de:
Especialista de Big Data

Director (a):
Ph.D., Omar Alexander León García

Fundación Universitaria Unipanamericana – Compensar
Facultad de Ingeniería, Ingeniería de (Especialización en Big Data)
Bogotá, Colombia
2023

Resumen

La violencia doméstica es un problema grave en Colombia, y el análisis de datos puede proporcionar información valiosa para comprender su naturaleza y magnitud. Este estudio se enfoca en diseñar un tablero de control que analice datos libres sobre la violencia doméstica en Colombia, con el objetivo de mejorar la comprensión de esta problemática. Se recopilarán datos relevantes de la base de datos pública de la Policía Nacional de Colombia y se desarrollará un modelo de análisis de datos para identificar patrones y tendencias en la violencia doméstica. El tablero de control permitirá visualizar la información de manera clara y accesible, facilitando la toma de decisiones informadas y el diseño de estrategias de prevención y atención. Sin embargo, esta investigación tiene limitaciones en cuanto a la disponibilidad y representatividad de los datos, así como en el análisis profundo de la problemática de la violencia doméstica.

Palabras clave: tablero de control, análisis de datos, violencia doméstica, policía nacional de Colombia.

Abstract

Domestic violence is a serious problem in Colombia, and data analysis can provide valuable insights into its nature and magnitude. This study focuses on designing a control dashboard that analyzes open data on domestic violence in Colombia, with the aim of improving understanding of this issue. Relevant data will be collected from the public database of the Colombian National Police, and a data analysis model will be developed to identify patterns and trends in domestic violence. The control dashboard will enable clear and accessible visualization of the information, facilitating informed decision-making and the design of prevention and intervention strategies. However, this research has limitations regarding data availability and representativeness, as well as a limited in-depth analysis of the domestic violence problem.

Keywords: Dashboard, data analysis, domestic violence, National Police of Colombia.

Contenido

PÁG.

RESUMEN ..... **V**

INTRODUCCIÓN **15**

ANTECEDENTES Y JUSTIFICACIÓN 17

PLANTEAMIENTO DEL PROBLEMA 19

OBJETIVO GENERAL 21

Objetivos Específicos *21*

ALCANCES Y LIMITACIONES 22

1. MARCO TEÓRICO **24**

1.1 MARCO CONCEPTUAL 24

1.1.1 Violencia Intrafamiliar *24*

1.1.2 Violencia Física *24*

1.1.3 Violencia Psicológica *25*

1.1.4 Violencia Sexual *25*

1.1.5 Violencia Económica *25*

1.2 MARCO REFERENCIA 26

1.3 MICROSOFT POWER BI 27

1.3.1 ¿Cuáles son las principales ventajas de Power BI? *27*

1.3.2 Tendencias del mercado en bi *30*

1.3.3 Análisis comparativos *30*

1.4. PYTHON 30

1.4.1 ¿Qué beneficios ofrece Python? *30*

1.4.2 ¿Cuáles son las características de Python? *31*

2. DISEÑO METODOLÓGICO **33**

3. RESULTADOS 53

4. DISCUSIÓN 59

5. CONCLUSIONES Y RECOMENDACIONES 61

5.1 CONCLUSIONES 61

5.2 RECOMENDACIONES 63

BIBLIOGRAFÍA 65

Listado de Figuras

Figura 1. Diseño de Arquitectura 33
Figura 2 . Librerías de Python 35
Figura 3 . Cargue del archivo .csv en Python e imprimir las 10 primeras líneas. 36
Figura 4. Listado general de los datos del Archivo .csv 36
Figura 5. Listado de las 5 primeras líneas de los datos del Archivo .csv 37
Figura 6. Listado los últimos 5 líneas del archivo .csv 37
Figura 7. Datos a nivel estadístico del Archivo .csv 37
Figura 8. Cantidad de filas y columnas del Archivo .csv 38
Figura 9. Las columnas de los datos del Archivo .csv 38
Figura 10. Listado general de los datos del Archivo .csv 38
Figura 11. Listado de tipo de datos de la columna Archivo .csv 39
Figura 12. Cantidad de valores nulos en cada columna del Archivo .csv 39
Figura 13. Listado los valores nulos del campo Genero en el Archivo .csv 40
Figura 14. Listado los valores nulos del campo Armas Medios en el Archivo .csv 40
Figura 15. Cambio de formato de fecha y visualizar en orden cronológico el campo de fecha del Archivo .csv 41
Figura 16. Cambio de formato del código Dane y Reemplazar no reportado por el signo - del Archivo .csv 41
Figura 17. Visualiza la información de los campos del Archivo .csv 42
Figura 18. Listado en orden numérico del Código Dane del Archivo .csv 42
Figura 19. Búsqueda de valores no reportado, reemplazar e imprimir la lista de genero del campo Genero del Archivo .csv 43
Figura 20. Listado en orden alfabético del Campo Género del Archivo .csv 43
Figura 21. Búsqueda de valores no reportado, reemplazar e imprimir la lista de genero del campo Armas Medios del Archivo .csv 43
Figura 22. Listado en orden alfabético del Campo armas medios del Archivo .csv 44
Figura 23. Reemplazo de valores nulos en el campo grupo Etario del Archivo .csv 44
Figura 24. Listado general de los datos del Archivo .csv 45
Figura 25. Listado en orden alfabético del campo grupo etario del Archivo .csv 45

Figura 26. Listado general de los datos del Archivo .csv 46
Figura 27. Generar Archivo .csv 46
Figura 28. Tablero de Control mediante Power Bi. 47
Figura 29. Mapa de Colombia visualiza la cantidad de casos por departamento. 47
Figura 30. Departamentos con mayor cantidad de casos 49
Figura 31. Cantidad de Casos por genero afectado 49
Figura 32. Porcentaje de armas usadas 50
Figura 33.Cantidad de casos por Año 51
Figura 34.Municipios con mayor cantidad de casos 51
Figura 35. Cantidad de Casos por grupo etario 52
Figura 36.Departamentos con mayor cantidad de casos. 53
Figura 37. Departamentos con mayor cantidad de casos. 54
Figura 38.Cantidad de Casos por genero afectado 55
Figura 39.Porcentaje de armas usadas 55
Figura 40.Cantidad de casos por Año 56
Figura 41.Municipios con mayor cantidad de casos 57
Figura 42. Cantidad de Casos por grupo etario 58
Figura 43 . Parte 1 de carta de la Evaluación del uso del Tablero de control 69
Figura 44. Parte 2 de carta de la Evaluación del uso del Tablero de control 70

Listado de Anexos

Anexo 1 69

Introducción

La violencia doméstica es un problema grave en Colombia y el análisis de datos puede ser una herramienta valiosa para comprender la naturaleza y la magnitud de esta problemática. Algunas de las problemáticas que se pueden estudiar mediante el análisis de datos sobre la violencia doméstica en Colombia. Estos individuos han experimentado diversas formas de discriminación, viendo vulnerados sus derechos y sufriendo maltrato físico o verbal, e incluso llegando a enfrentar desenlaces violentos en algunos casos (Martha Baez,2011).

En los últimos tiempos, la ciudad de Bogotá ha sido testigo de una preocupante serie de hechos violentos que han afectado a diversos grupos de la población, especialmente a mujeres, adultos mayores y menores de edad. Estos lamentables sucesos han despertado la atención de la sociedad y de las autoridades encargadas de proteger los derechos de las personas más vulnerables en el ámbito familiar (Instituto Colombiano de Bienestar Familiar,2013).

Se desarrollaron las diferentes sesiones de trabajo y un cronograma de estas para ir definiendo paso a paso la problemática a analizar en el trabajo de grado y mediante la asignación de tareas, ideas y habilidades del grupo hemos ido alineando a las expectativas de las actividades y del documento final.

Es por ello, que comenzaremos por definir el siguiente trabajo de grado con el diagnóstico inicial al que nos invita la actividad de reconocimiento y que presentamos a continuación y así poder concretar aspectos como:

La investigación tendrá un enfoque cualitativo y el insumo que utilizaremos proviene del banco de datos públicos indicado y validado por el docente. Los cuestionamientos son diversos dado que no se tiene una hipótesis definida, sino que se encontrarán factores o comportamientos que aún se esperan formalizar. El caso que analizaremos es una data sobre

hechos reales, por lo tanto, se sugerirán análisis relacionados con la problemática social, económica y política del país, de acuerdo con la data histórica y hasta el último periodo actualizado de la base de datos abiertos. El desarrollo del análisis se visualizará mediante un tablero de control para el análisis de la violencia intrafamiliar en Colombia desde el año 2010 al 2023.

En conclusión, el análisis de datos sobre la violencia doméstica en Colombia puede ayudar a comprender mejor esta problemática y a tomar medidas para prevenirla y abordarla.

Antecedentes y Justificación

La violencia domestica o más conocida también como intrafamiliar viene desde tiempos atrás, en donde se muestra un problema de abuso del esposo hacia la esposa y en la falta de mejoras de la vida de una persona de la tercera edad. Estas personas han tenido que vivir de muchas formas de discriminaciones; vulnerando sus derechos y generando maltratos físicos o verbales e incluso sus fallecimientos de formas violentas (Martha Baez,2011).

En los últimos tiempos se ha descubierto en la ciudad de Bogotá varios hechos violentos de que han sido víctimas mujeres, adultos mayores y menos de edad. Los esfuerzos de las campañas de concientizar a la población por parte de los entes de control encargadas cuidar los derechos de las personas más vulnerables dentro un grupo familiar, las comisarías de familia reciben y atienden esta problemática dando en lo mejor posible soluciones y recomendaciones para finalizar los conflictos, cuando hay lugar a ellos (Instituto Colombiano de Bienestar Familiar,2013).

Este tipo de violencia se evidencia en patrones culturales en donde las relaciones familiares están respaldadas en el tema de la Sumisión o dominación, que genera una gran preeminencia al poder masculino, cuya representación para solucionas los problemas es a través de la Violencia (Personería de Bogotá,2003).

La investigación sobre la violencia doméstica y su análisis a través de un tablero de control es importante por varias razones:

- Contribuye a la prevención y atención de la violencia doméstica: Un tablero de control que permita analizar y visualizar los datos sobre la violencia doméstica en Colombia puede ayudar a identificar patrones, tendencias y áreas de enfoque para prevenir y atender la violencia doméstica.

- Fomenta el uso de datos abiertos: La investigación también contribuye al fomento del uso de datos abiertos, ya que se basa en el análisis de datos públicos disponibles de manera gratuita. Esto puede ayudar a fomentar la transparencia y la rendición de cuentas en el manejo de la información pública en Colombia.

▪ Facilita la toma de decisiones informadas: La visualización de los datos sobre la violencia doméstica en un tablero de control puede facilitar la toma de decisiones informadas por parte de los tomadores de decisiones, ya que pueden analizar los datos de manera más clara y rápida, lo que puede influir en la implementación de políticas públicas y estrategias para prevenir y atender la violencia doméstica en Colombia.

▪ Sirve como una herramienta de sensibilización: El tablero de control también puede ser utilizado como una herramienta de sensibilización para la población general, ya que puede ayudar a aumentar la conciencia sobre la violencia doméstica y sus impactos en la sociedad.

Planteamiento del Problema

Según una investigación realizada por el Instituto Nacional de Medicina Legal y Ciencias Forenses de Colombia, en el año 2021 se registraron 52.019 casos de violencia intrafamiliar en el país (Instituto Nacional de Medicina Legal y Ciencias Forenses, 2021).

Según Rodríguez y Zuleta (2020), la violencia doméstica en Colombia es un problema complejo que afecta a una gran cantidad de personas en todo el país. Los autores señalan que, si bien existen algunas estadísticas disponibles sobre este tema, aún hay muchas lagunas en el conocimiento que dificultan las formulaciones efectivas para prevenir y abordar la violencia doméstica.

Por su parte, Oviedo y Zapata (2021) argumentan que el análisis de datos es una herramienta esencial para entender la violencia doméstica en Colombia. Los autores destacan la importancia de recopilar información detallada sobre las víctimas, los perpetradores y las circunstancias en las que se produce la violencia. También enfatizan la necesidad de realizar un análisis exhaustivo de los datos recopilados para identificar patrones y tendencias que puedan ayudar a comprender mejor este fenómeno.

Algunas de las problemáticas que se pueden estudiar mediante el análisis de datos sobre la violencia doméstica en Colombia en el 2023 podrían incluir: Identificación de patrones y tendencias en la incidencia de la violencia doméstica en Colombia a lo largo del tiempo y en diferentes regiones del país (Ruiz-Gómez, 2021). Análisis de los factores de riesgo asociados a la violencia doméstica en Colombia, tales como el género, la edad, la educación y el nivel socioeconómico de las víctimas y los agresores (Gómez, 2020).

En general, el análisis de datos sobre la violencia doméstica en Colombia puede proporcionar información valiosa para entender y abordar este grave problema social.

Teniendo en cuenta lo anteriormente expuesto, se ve la necesidad de plantear el siguiente interrogante para esta investigación:

¿Cómo puede construirse un tablero de control que analice datos públicos sobre la violencia doméstica en Colombia y qué información relevante puede obtenerse a partir de él para analizar esta problemática?

Objetivo General

Construir un tablero de control para el análisis de datos públicos sobre la violencia doméstica en Colombia, que permita identificar áreas prioritarias de intervención y desarrollar estrategias efectivas de prevención.

Objetivos Específicos

- Identificar y recopilar datos relevantes sobre la violencia doméstica en Colombia, proveniente de la base de datos Publica de la Policía Nacional de Colombia.
- Desarrollar un modelo de análisis de datos que permita identificar patrones y tendencias en la violencia doméstica en el país, considerando factores como el tipo de violencia, la zona geográfica, el perfil de los agresores y las víctimas, entre otros.
- Diseñar el tablero de control, utilizando herramientas de visualización de datos para presentar la información de manera clara y accesible, y que permita una navegación intuitiva.
- Evaluar la efectividad del tablero de control en la identificación de patrones y tendencias en la violencia doméstica en Colombia, y su capacidad para generar información útil para el diseño de estrategias de prevención y atención.

Alcances y Limitaciones

Dado que el objetivo de esta investigación es diseñar un tablero de control que permita analizar datos libres sobre la violencia doméstica en Colombia, para mejorar la comprensión de la problemática. Para ello, se establecerá un alcance dentro de los siguientes términos:

▪ Fuentes de datos: se utilizarán datos libres disponibles públicamente sobre la violencia doméstica en Colombia, provenientes de la policía nacional (policia.gov.co,2023).

▪ Análisis de datos: se realizará un análisis básico de los datos para su posterior visualización en el tablero de control, sin profundizar en un análisis de fondo sobre la problemática de la violencia doméstica.

▪ Tablero de control: se desarrollará un tablero de control que permita la visualización interactiva de los datos analizados, utilizando la herramienta de visualización de datos Power BI.

▪ Funcionalidades del tablero de control: el tablero de control permitirá la visualización de la evolución temporal de la violencia doméstica, la identificación de patrones y tendencias según variables como la zona geográfica, el tipo de violencia, el perfil de los agresores y las víctimas, entre otros.

▪ El alcance del proyecto abarcaría el diseño y construcción de un tablero de control que analice datos públicos sobre la violencia doméstica en Colombia. Esto implica la recopilación de datos (Públicos) relevantes de fuentes confiables y su posterior procesamiento y visualización en el tablero.

Por otro lado, las principales limitantes en la investigación son:

▪ Limitaciones en los datos: La investigación se basa en la utilización de datos libres disponibles públicamente, lo que puede limitar la cantidad y calidad de la información disponible.

- Limitaciones en el análisis de los datos: Dado que el objetivo principal de la investigación es el diseño y desarrollo del tablero de control, el análisis de los datos se limitará a un nivel básico, lo que puede reducir la profundidad del análisis.

- Limitaciones en la disponibilidad de herramientas: El desarrollo del tablero de control puede estar limitado por la disponibilidad de herramientas de visualización de datos y programación.

- Limitaciones en la representatividad de los datos: La selección de fuentes de datos públicas puede limitar la representatividad de los datos, ya que no todos los casos de violencia doméstica pueden ser reportados o registrados.

Dado la naturaleza de la especialización cursada, el enfoque principal que se dará en esta investigación será en el manejo de la data, hallazgo de patrones, revisión de variables y comportamientos, deducciones hipotéticas sobre el comportamiento de estas y finalmente una visualización de los aspectos a resaltar que sean resultado del análisis respectivo.

Sobre las distintas variables será necesario realizar una contextualización en los ámbitos sociales y políticos principalmente que nos darán la pauta para generar criterios de visualización en el cuadro de mando que estaremos construyendo.

Para Colombia específicamente podríamos relacionar como en la mayoría de Latinoamérica el maltrato intrafamiliar con escases a nivel económico e intelectual, sin embargo, sería interesante revisar la data para encontrar si existes otros tipos de segmentaciones en la ejecución de estos actos, no solo por edad o por género, sino en el cruce de estas mismas variables reconocer otro tipo de patrones que puedan estar escondiéndose tras las causas cliché.

1. Marco teórico

1.1 Marco conceptual

En el documento se va a tratar del tema de los antecedentes de violencia doméstica en Colombia, en el cual se debe comenzar por saber que la violencia se conoce como la interferencia física que ejerce un individuo o un grupo en el cuerpo de un tercero, sin su consentimiento, cuyas consecuencias pueden ir desde una conmoción, una contusión o un rasguño, una inflamación o un dolor de cabeza o a un hueso roto, un ataque al corazón, la pérdida de un miembro e incluso la muerte (Jhon Keane,2000).

1.1.1 Violencia Intrafamiliar

La violencia intrafamiliar o violencia domestica puede definirse como una situación de abuso de poder o maltrato físico o psicológico, de un miembro de la familia, sobre todo. Puede manifestarse a través de golpes e incidentes graves, como también insultos, manejo económico, amenazas, chantajes, control de las actividades, abuso sexual, aislamiento de familiares y amistades, prohibición a trabajar fuera de la casa, abandono afectivo, humillaciones o no respetar las opiniones.

Estos son solo algunos de las situaciones que podemos citar como ejemplos de lo que es la violencia domestica o violencia intrafamiliar que en nuestro país está muy marcada por el hecho de que se dé el machismo en muchos lugares principalmente rurales y pobres de nuestra república (Ramírez, 2021).

1.1.2 Violencia Física

La violencia física son actos que atentan o agraden el cuerpo de la persona tales como empujones, bofetadas, golpes de puño, golpes de pies, etc. Según las estadísticas, la mayoría de las personas han sufrido alguna vez en su vida violencia de este tipo; violencia física, la cual se hace presente en todos los lugares, tanto que se nos ha hecho natural ver a dos personas golpeándose o se ha visto normal que un padre eduque a su hijo a través de los golpes y los insultos. Actualmente todos los tipos de violencia están

penados por la ley y son acreedores a multas o cárcel a las personas que ejercen y promueven la violencia doméstica (Ramírez, 2021).

1.1.3 Violencia Psicológica

La violencia psicológica son actitudes que tienen como objetivo causar el temor, intimidar y controlar las conductas, sentimientos y pensamientos de la persona a quien se está agrediendo como las descalificaciones, insulto, control, etc. Según lo que entendemos por violencia intrafamiliar son golpes, maltratos, lagrimas, alcohol o incluso hasta la muerte, pero pocas veces tomamos en cuenta que otro tipo de violencia intrafamiliar es <<La violencia psicológica>>, que se hace presente sin ningún maltrato de tipo físico, pero que tiene las mimas o peores repercusiones en el individuo maltratado, puesto que crea seres inseguros e incapaces de resolver sus problemas, gente aislada que se niega a recibir ayuda profesional, gente cerrada a su ideología de los roles de género, etc (Ramírez, 2021).

1.1.4 Violencia Sexual

La violencia sexual se define como la imposición de actos de carácter sexual contra la voluntad de la otra persona. Como por ejemplos exposición a actividades sexuales no deseadas, la manipulación a través de la sexualidad, tocamientos, miradas, caricias que quien las recibe no las desea, violación; es decir la consumación del acto sexual por la fuerza. Este tipo de violencia se da muchas veces entre los parientes de los niños como tíos, primos, abuelos, etc. Que obligan a los niños a realizar actos sexuales frente a ellos, tales como masturbación o mostrarles sus órganos sexuales, etc (Ramírez, 2021).

1.1.5 Violencia Económica

La violencia económica se refiere a no cubrir las necesidades básicas de la persona y ejercer control a través de recursos económicos. Este tipo de violencia es muy frecuentada sobre las personas de la tercera edad, aquellas que no pueden sostener su economía por si solos y que necesitan de alguien que solvente sus gastos, puesto que en muchos de los casos la persona que aporta el dinero también es quien limita al adulto mayor sobre sus gastos y no cubre con

las necesidades básicas de este. Las repercusiones de este tipo de violencia también son psicológicas ya que puede sentirse inútil e incapaz de remediar la situación y de estar sometido a alguien y sentirse carga de este lo cual viene terminando en huida del hogar o suicidio (Ramírez, 2021).

1.2 Marco Referencia

En los últimos años, ha existido un tremendo incremento de las capacidades de almacenamiento y generación de bases de datos en todas las áreas de conocimiento, incluida la criminología. El Estado recolecta información sobre todos los ciudadanos, manejada por los departamentos de policía, sistemas judiciales y organizaciones de desarrollo social. La agregación de esta información administrativa permite tener antecedentes en distintas dimensiones para todas las interacciones entre un individuo y el aparato estatal. Esto abre posibilidades insospechadas en cuanto al análisis de datos, que no existían hasta hace algunos años (Ridgeway, Policing in the Era of Big Data, 2018).

Los gobiernos en distintos países del mundo están abordando el problema de la reincidencia delictual mediante el análisis de datos administrativos y el uso de modelos matemáticos, para comprender los patrones que pueden explicarla. Ozkan (2017) realizó un estudio sobre la aplicación de técnicas de aprendizaje automatizado para la predicción de reincidencia. Los resultados revelan que en Estados Unidos la tasa de reincidencia en promedio es de un 70% de los individuos que han sido liberados dentro de una ventana de tres años. Esta tasa varía según el tipo de delitos, siendo más alta para aquellos que afectan a la propiedad. Otros ejemplos del uso de aprendizaje automatizado han sido presentados por Bartolucci et al. (2007) y Zeng et al. (2017), quienes desarrollaron métodos matemáticos para la predicción de trayectorias y reincidencia criminal. A pesar del beneficio que este tipo de metodologías pueda traer en este contexto, también existen críticas sobre las limitaciones de los modelos predictivos y la extensión de las conclusiones que se puedan sacar a partir de estas herramientas de big data (Dressel, 2018).

1.3 Microsoft Power BI

Power BI es una plataforma unificada y escalable de inteligencia empresarial (BI) con funciones de autoservicio apta para grandes empresas. Conéctese a los datos, visualícelos e incorpore sin problemas objetos visuales en las aplicaciones que usa todos los días (Microsoft, 2023). Según Microsoft power BI es un conjunto de aplicaciones de análisis de negocios que permite analizar datos y compartir información. Los paneles de Power BI ofrecen a los usuarios una vista de 360 grados con sus métricas más importantes en un mismo lugar. La información se actualiza en tiempo real y está disponible en todos sus dispositivos.

1.3.1 ¿Cuáles son las principales ventajas de Power BI?

Debido a sus funcionalidades, el uso de esta herramienta ofrece grandes beneficios. En esta publicación se detallan los principales (dqsconsulting, 2021).

- **Coste bajo**

Power BI tiene dos versiones: Power BI Pro y Power BI Premium. El nivel Power Bi PRO, que sería el de entrada, ya ofrece todas las funcionalidades del servicio con un coste bajo al mes. Hay que recordar, además, que es una herramienta que forma parte del paquete Office. La versión Power BI Premium, aunque tiene un coste más elevado, también ofrece más ventajas: permite utilizarse por cientos de usuarios a la vez y actualizar de manera ágil los datos más de 8 veces al día. Para escoger entre una opción u otra se debe tener en cuenta el número de licencias que se necesitarán; es decir, las personas que deberán tener acceso (dqsconsulting, 2021).

- **Permite hacer múltiples análisis de datos en un solo informe o panel**

Con una única visualización, y presentado de manera visual y atractiva, se podrán consultar en un mismo dashboard múltiples análisis a la vez. Es una manera de presentar numerosos datos en un mismo documento y que, además, sean fáciles de comprender (dqsconsulting, 2021).

- **Los paneles se pueden compartir**

Es una herramienta colaborativa, permite que estén conectados a ella un alto número de usuarios y, además, que los paneles e informes se puedan publicar o compartir con toda la compañía. Esto facilita que se fomente la cultura de trabajo en equipo y el análisis de datos

dentro de la organización. Es decir, potencia el trabajo en equipo y la colaboración entre compañeros, partners y clientes. Al estar alojado en la nube, solo se necesita tener conexión a Internet para poder acceder a toda la información (dqsconsulting, 2021).

- **Es una herramienta fácil de utilizar**

  Power BI es una herramienta muy intuitiva y, por lo tanto, fácil de usar para usuarios de todos los niveles. Se pueden crear informes y paneles sin tener conocimientos técnicos, especialmente en la versión de ordenador. En cualquier caso, si se quiere hacer un uso más profesional de la herramienta y aprovechar todo su potencial, sí que será necesaria una formación (dqsconsulting, 2021).

- **Incluye las herramientas Power View y Power Map**

  En Power BI están incorporadas las herramientas de visualización de imágenes y gráficos en 3D Power View y Power Map, que permiten también ver datos en mapas geográficas. Estas herramientas serán muy útiles dependiendo de los datos que se quieran mostrar en el panel o en el informe, ya que permiten un análisis avanzado y mostrar, por ejemplo, los datos según variables de ubicación o contexto (dqsconsulting, 2021).

- **El uso de cuadros de mando**

  El resultado final donde se presentan los datos son paneles, o cuadros de mando muy intuitivos que generan una idea muy visual de la situación de la empresa o de los datos que se quieran mostrar. Los gráficos que utiliza son muy interactivos y, al pasar el ratón por encima, o al hacer clic, puede aparecer más información detallada. Además, da la opción de filtrar los datos seleccionado fecha, clientes, países, etcétera. También cabe destacar que es totalmente personalizable. El usuario puede escoger las formas que desee, al igual que los colores y la distribución final del cuadro de mando (dqsconsulting, 2021).

- **Datos actualizados**

  No todas las herramientas de gestión de datos permiten consultas en vivo ni garantizan que se esté viendo la última versión del documento. Esto no pasa con Power BI, ya que al ser colaborativo o estar alojado en la nube, los datos que se mostrarán serán siempre

los más recientes. Con la versión Power BI Pro se puede actualizar la caché de datos hasta 8 veces al día y, con Power BI Premium, hasta 48 (dqsconsulting, 2021).

- **Se integra a la perfección con Microsoft Excel**

Esta es una de sus mayores ventajas, ya que la mayoría de las empresas están acostumbradas a trabajar con Excel y tienen en este programa la mayoría de sus informes. Pasar de Excel a Power BI es un gran cambio, aunque sencillo, y el hecho de que sean dos herramientas integradas facilita y mucho cualquier proceso. Se pueden exportar y conectar de una manera sencilla los paneles de Power BI con los datos almacenados en Excel (dqsconsulting, 2021).

- **Tiene una interfaz de lenguaje natural**

El hecho de que tenga una interfaz de lenguaje natural implica que se puedan realizar consultas de una manera muy fácil, de la misma manera en que se le haría la pregunta a una persona. La herramienta responde con la información solicitada y la puede presentar mediante un dato, una tabla o un gráfico. Una forma muy sencilla y rápida de tener respuesta a todas las dudas que impliquen datos (dqsconsulting, 2021).

- **Es una herramienta de Inteligencia Artificial**

Al tratarse de una herramienta de Inteligencia Artificial, Power BI analiza los datos actuales y del pasado que se le faciliten, pero, además, también permite detectar tendencias y hacer predicciones (forecating) de lo que sucederá en el futuro en el negocio sobre actividades de ventas. Esto lo realiza partiendo de una base de datos que se le deben facilitar (dqsconsulting, 2021).

- **Datos accesibles desde cualquier sistema**

Como ya se ha explicado con anterioridad, se puede acceder a los paneles de Power BI desde cualquier dispositivo: ya sea ordenador, tablet o móvil. También desde cualquier sistema, Windows, iOS o Android, a cualquier hora y en cualquier lugar. Lo único que se necesita para acceder a esta herramienta es tener conexión a Internet (dqsconsulting, 2021).

1.3.2 Tendencias del mercado en bi

El software de Business Intelligence (BI) es una de las tecnologías de software más populares en el mercado hoy en día. De hecho, en una encuesta a más de 10 mil ejecutivos de tecnología, el 23.5% respondió que una plataforma de BI o BA (análisis de negocios) es su próxima compra de software planificada. Elegir un proveedor de software de BI es una tarea que se debe considerar seriamente. Esta comparación de software de inteligencia empresarial está destinada a ser utilizada como punto de partida para guiar a los profesionales de TI encargados de tomar esta decisión (Mamani, 2018).

1.3.3 Análisis comparativos

Es habitual la selección de una herramienta siempre va a depender de las necesidades de la empresa, el perfil de los usuarios y entre otras consideraciones. La más común es el análisis de datos y generación de informes a través de gráficos estadísticos que es un valor que ayuda a mejorar la toma de decisiones en diferentes ámbitos. Cuando los profesionales de TI inician la selección de alguna de herramienta plantean criterios como, por ejemplo: por precio, autoservicio, rendimiento en la nube, por tener ETL (Extract, Transform and Load), conectividad, soporte big data, visualizaciones amigables y experiencia de usuario (Mamani, 2018).

1.4. Python

Python es un lenguaje de programación ampliamente utilizado en las aplicaciones web, el desarrollo de software, la ciencia de datos y el machine learning (ML). Los desarrolladores utilizan Python porque es eficiente y fácil de aprender, además de que se puede ejecutar en muchas plataformas diferentes. El software Python se puede descargar gratis, se integra bien a todos los tipos de sistemas y aumenta la velocidad del desarrollo (aws.amazon.com, 2023).

1.4.1 ¿Qué beneficios ofrece Python?

Los beneficios de Python incluyen los siguientes:

- Los desarrolladores pueden leer y comprender fácilmente los programas de Python debido a su sintaxis básica similar a la del inglés.
- Python permite que los desarrolladores sean más productivos, ya que pueden escribir un programa de Python con menos líneas de código en comparación con muchos otros lenguajes.
- Python cuenta con una gran biblioteca estándar que contiene códigos reutilizables para casi cualquier tarea. De esta manera, los desarrolladores no tienen que escribir el código desde cero.
- Los desarrolladores pueden utilizar Python fácilmente con otros lenguajes de programación conocidos, como Java, C y C++.
- La comunidad activa de Python incluye millones de desarrolladores alrededor del mundo que prestan su apoyo. Si se presenta un problema, puede obtener soporte rápido de la comunidad.
- Hay muchos recursos útiles disponibles en Internet si desea aprender Python. Por ejemplo, puede encontrar con facilidad videos, tutoriales, documentación y guías para desarrolladores.
- Python se puede trasladar a través de diferentes sistemas operativos de computadora, como Windows, macOS, Linux y Unix (aws.amazon.com, 2023).

1.4.2 ¿Cuáles son las características de Python?

Las características siguientes del lenguaje de programación Python lo hacen único:

- **Un lenguaje interpretado**

Python es un lenguaje interpretado, lo que significa que ejecuta directamente el código, línea por línea. Si existen errores en el código del programa, su ejecución se detiene. Así, los programadores pueden encontrar errores en el código con rapidez.

- **Un lenguaje fácil de utilizar**

Python utiliza palabras similares a las del inglés. A diferencia de otros lenguajes de programación, Python no utiliza llaves. En su lugar, utiliza sangría.

- **Un lenguaje tipeado dinámicamente**

Los programadores no tienen que anunciar tipos de variables cuando escriben código porque Python los determina en el tiempo de ejecución. Debido a esto, es posible escribir programas de Python con mayor rapidez.

- **Un lenguaje de alto nivel**

Python es más cercano a los idiomas humanos que otros lenguajes de programación. Por lo tanto, los programadores no deben preocuparse de sus funcionalidades subyacentes, como la arquitectura y la administración de la memoria.

- **Un lenguaje orientado a los objetos**

Python considera todo como un objeto, pero también admite otros tipos de programación, como la programación estructurada y la funcional (aws.amazon.com, 2023).

2. Diseño Metodológico

Se diseño este modelo con el objetivo de analizar la violencia intrafamiliar en Colombia que permita mediante la limpieza y extracción de datos, carga de los datos para luego ser visualizado en un tablero de control en Power BI. Para ello, se utilizará información de fuentes secundarias, como bases de datos de la policía de nacional de Colombia (policia.gov.co,2023).

Para iniciar se va a utilizar el proceso de extracción, transformación y limpieza de los datos mediante la herramienta de Python que nos ayuda detectar y corregir o eliminar los registros corruptos o incorrectos de un conjunto de datos, tabla o base de datos, para luego ser entregados a una herramienta de llamada Power bi para si visualización Dare Planet Technology (2022).

A continuación, se describe el proceso que se ha utilizado para la implementación del proyecto:

Análisis del requerimiento: Se identifica la información que se obtiene de los datos públicos de violencia doméstica en Colombia entre los periodos del 2010 al marzo 2023.
Diseño de arquitectura: Se genera un diseño donde se visualiza los procesos a realizar sobre el requerimiento del proyecto.



Figura 1. Diseño de Arquitectura
Fuente: elaboración propia

- **Fuente de datos:** La fuente de datos origen se encuentra alojado en el sitio Web de la policía Nacional en un formato csv Excel y con proceso de recopilación y unificación de los datos históricos 2010 al marzo 2023.

- **Proceso Extracción, Transformación y limpieza:** El proceso de Extracción, Transformación y limpieza se utiliza la herramienta de Python.

- **Visualización y Análisis: Mediante** la aplicación de Power BI desktop se puede cargar los datos, los cuales se generan los diferentes gráficos para luego analizarlos.

Integración de datos: Se realiza los pasos de extracción, transformación y limpieza de datos utilizando la herramienta de Python.

Diseño del Dashboard con la aplicación de Power BI: Se genera un diseño donde se puede analizar los departamentos, genero, armas utilizadas, grupo etario y en qué año generaron violencia doméstica Colombia.

Validación del tablero: El objetivo de esta etapa es evaluar la calidad, pertinencia y aplicabilidad del tablero y los resultados obtenidos mediante la revisión y retroalimentación de un experto temático.

Python

Python es un lenguaje de programación ampliamente utilizado en las aplicaciones web, el desarrollo de software, la ciencia de datos y el machine learning (ML). Los desarrolladores utilizan Python porque es eficiente y fácil de aprender, además de que se puede ejecutar en muchas plataformas diferentes (Amazon Web Services,2023).

Pasos para limpieza y extracción de datos

- Exploración
- Tipo de datos
- Tratar los datos perdidos
- Valores no permitidos
- Manejo de fechas
- Valores duplicados

Power BI

Power BI es una plataforma unificada y escalable de inteligencia empresarial (BI) con funciones de autoservicio apta para grandes empresas. Conectando a los datos, visualizando e incorporando sin problemas con objetos visuales en las aplicaciones que se usa todos los días (Microsoft,2023).

Integración de datos

Proceso de Limpieza de Datos mediante la aplicación de Python

En la figura 2, se realiza la importación de las librerías que se requieren para la compilación del código fuente en Python.

```
#importar libreria
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import datetime as dt
import seaborn as sns
```

Figura 2 . Librerías de Python
Fuente: elaboración propia

En la figura 3, Se realiza la carga de datos del Reporte en formato .csv y se visualiza las primeras 10 líneas del archivo.

```
##Cargar el conjunto de datos

datos = pd.read_csv('ReporteViolenciadomesticaPNC.csv',low_memory=False)
datos.head(10) # Imprimir los primeras 10 lineas
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
0	ATLÁNTICO	BARRANQUILLA (CT)	8001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	MASCULINO	ADULTOS	1
1	BOYACÁ	DUITAMA	15238000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
2	CAQUETÁ	PUERTO RICO	18592000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	MASCULINO	ADULTOS	1
3	CASANARE	MANÍ	85139000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
4	CUNDINAMARCA	BOGOTÁ D.C. (CT)	11001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
5	SUCRE	SINCELEJO (CT)	70001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
6	VALLE	CALI (CT)	76001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	2
7	VALLE	CALI (CT)	76001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	MASCULINO	ADULTOS	2
8	VALLE	TULUÁ	76834000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
9	HUILA	PITALITO	41551000	ARMA DE FUEGO	1/01/2010	FEMENINO	ADULTOS	1

Figura 3 . Cargue del archivo .csv en Python e imprimir las 10 primeras líneas.
Fuente: elaboración propia

En la figura 4, Se muestra el listado general de los datos y nos indica el número de filas con las columnas.

```
#Vistazo general al dataFrame
datos
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
0	ATLÁNTICO	BARRANQUILLA (CT)	8001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	MASCULINO	ADULTOS	1
1	BOYACÁ	DUITAMA	15238000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
2	CAQUETÁ	PUERTO RICO	18592000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	MASCULINO	ADULTOS	1
3	CASANARE	MANÍ	85139000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
4	CUNDINAMARCA	BOGOTÁ D.C. (CT)	11001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
...	...	...	...	...	...	...	...	...
548037	VICHADA	Puerto Carreño (CT)	99001000	SIN EMPLEO DE ARMAS	21/02/2023	FEMENINO	MENORES	1
548038	VICHADA	Puerto Carreño (CT)	99001000	SIN EMPLEO DE ARMAS	26/02/2023	FEMENINO	ADULTOS	1
548039	VICHADA	Puerto Carreño (CT)	99001000	SIN EMPLEO DE ARMAS	02/03/2023	FEMENINO	ADOLESCENTES	1
548040	VICHADA	Puerto Carreño (CT)	99001000	SIN EMPLEO DE ARMAS	14/03/2023	FEMENINO	ADULTOS	1
548041	VICHADA	Santa Rosalía	99624000	SIN EMPLEO DE ARMAS	31/03/2023	FEMENINO	ADOLESCENTES	1

548042 rows × 8 columns

Figura 4. Listado general de los datos del Archivo .csv
Fuente: elaboración propia

En la figura 5, Se muestra los datos iniciales del reporte de violencia doméstica.

```
#Visualizar los inicial datos del dataframe
datos.head()
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
0	ATLÁNTICO	BARRANQUILLA (CT)	8001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	MASCULINO	ADULTOS	1
1	BOYACÁ	DUITAMA	15238000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
2	CAQUETÁ	PUERTO RICO	18592000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	MASCULINO	ADULTOS	1
3	CASANARE	MANÍ	85139000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1
4	CUNDINAMARCA	BOGOTÁ D.C. (CT)	11001000	ARMA BLANCA / CORTOPUNZANTE	1/01/2010	FEMENINO	ADULTOS	1

Figura 5. Listado de las 5 primeras líneas de los datos del Archivo .csv
Fuente: elaboración propia

En la figura 6, Se muestra los últimos datos del reporte de violencia doméstica.

```
#Visualizar los ultimos datos del dataframe

datos.tail()
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
548037	VICHADA	Puerto Carreño (CT)	99001000	SIN EMPLEO DE ARMAS	21/02/2023	FEMENINO	MENORES	1
548038	VICHADA	Puerto Carreño (CT)	99001000	SIN EMPLEO DE ARMAS	26/02/2023	FEMENINO	ADULTOS	1
548039	VICHADA	Puerto Carreño (CT)	99001000	SIN EMPLEO DE ARMAS	02/03/2023	FEMENINO	ADOLESCENTES	1
548040	VICHADA	Puerto Carreño (CT)	99001000	SIN EMPLEO DE ARMAS	14/03/2023	FEMENINO	ADULTOS	1
548041	VICHADA	Santa Rosalía	99624000	SIN EMPLEO DE ARMAS	31/03/2023	FEMENINO	ADOLESCENTES	1

Figura 6. Listado los últimos 5 líneas del archivo .csv
Fuente: elaboración propia

En la figura 7, Se muestra los datos a nivel estadístico como el total de líneas, media, media estándar, mínimos, percentiles y el máximo.

```
#describe los datos nivel estadistico

datos.describe().T
```

	count	mean	std	min	25%	50%	75%	max
CANTIDAD	548042.0	1.785228	3.847218	1.0	1.0	1.0	1.0	143.0

Figura 7. Datos a nivel estadístico del Archivo .csv
Fuente: elaboración propia

En la figura 8, Se muestra el tamaño del dataframe o sea la cantidad de líneas y columnas que conforma el reporte de violencia doméstica, con 8 columnas o variables de información:

```
#consultar el tamaño del dataframe
datos.shape
```

```
(548042, 8)
```

Figura 8. Cantidad de filas y columnas del Archivo .csv
Fuente: elaboración propia

En la figura 9, Se muestra las columnas del reporte de la violencia doméstica:

```
# Las columnas del dataframe

datos.columns
```

```
Index(['DEPARTAMENTO', 'MUNICIPIO', 'CODIGO DANE', 'ARMAS MEDIOS', 'FECHA',
       'GENERO', 'GRUPO ETARIO', 'CANTIDAD'],
      dtype='object')
```

Figura 9. Las columnas de los datos del Archivo .csv
Fuente: elaboración propia

En la figura 10, Se muestra tipos de formatos de cada de columna:

```
In [130]: # tipos datos del dataframe
          datos.dtypes

Out[130]: DEPARTAMENTO    object
          MUNICIPIO       object
          CODIGO DANE     object
          ARMAS MEDIOS    object
          FECHA           object
          GENERO          object
          GRUPO ETARIO    object
          CANTIDAD         int64
          dtype: object
```

Figura 10. Listado general de los datos del Archivo .csv
Fuente: elaboración propia

En la figura 11, Se muestra la información de los tipos de datos del Dataframe o sea de cada columna del reporte de violencia doméstica:

```
#Información de tipos de dato del dataFrame
datos.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 548042 entries, 0 to 548041
Data columns (total 8 columns):
 #   Column         Non-Null Count   Dtype 
---  ------         --------------   ----- 
 0   DEPARTAMENTO   548042 non-null  object
 1   MUNICIPIO      548042 non-null  object
 2   CODIGO DANE    548042 non-null  object
 3   ARMAS MEDIOS   548040 non-null  object
 4   FECHA          548042 non-null  object
 5   GENERO         548039 non-null  object
 6   GRUPO ETARIO   546427 non-null  object
 7   CANTIDAD       548042 non-null  int64 
dtypes: int64(1), object(7)
memory usage: 33.4+ MB
```

Figura 11. Listado de tipo de datos de la columna Archivo .csv
Fuente: elaboración propia

En la figura 12, Se muestra la cantidad de valores nulos en las columnas del reporte de violencia Domestica:

```
#Localizar valores Nulos
datos.isnull().sum()
```

```
DEPARTAMENTO       0
MUNICIPIO          0
CODIGO DANE        0
ARMAS MEDIOS       2
FECHA              0
GENERO             3
GRUPO ETARIO    1615
CANTIDAD           0
dtype: int64
```

Figura 12. Cantidad de valores nulos en cada columna del Archivo .csv
Fuente: elaboración propia

En la figura 13, se selecciona los valores en error N nulos del campo Genero del reporte de violencia doméstica:

```
#seleccionar filas concretas  con tiene valores NAN

fila_con_nan = datos.isnull()['GENERO']
filas_con_nan = datos[fila_con_nan]
print(filas_con_nan)
```

```
        DEPARTAMENTO         MUNICIPIO CODIGO DANE         ARMAS MEDIOS  \
469436  CUNDINAMARCA  BOGOTÁ D.C. (CT)    11001000  SIN EMPLEO DE ARMAS
474304     RISARALDA      PEREIRA (CT)    66001000  SIN EMPLEO DE ARMAS
476771         VALLE           PALMIRA    76520000  SIN EMPLEO DE ARMAS

             FECHA GENERO GRUPO ETARIO  CANTIDAD
469436   1/03/2021    NaN          NaN         1
474304   1/03/2021    NaN          NaN         1
476771  26/03/2021    NaN          NaN         1
```

Figura 13. Listado los valores nulos del campo Genero en el Archivo .csv
Fuente: elaboración propia

En la figura 14, se selecciona los valores en error o nulos del campo Armas Medios del reporte de violencia doméstica:

```
#seleccionar filas concretas  con tiene valores NAN

fila_con_nan = datos.isnull()['ARMAS MEDIOS']
filas_con_nan = datos[fila_con_nan]
print(filas_con_nan)
```

```
        DEPARTAMENTO  MUNICIPIO CODIGO DANE ARMAS MEDIOS  FECHA    GENERO  \
476968  CUNDINAMARCA      TAUSA    25793000          NaN  44294  FEMENINO
476969  CUNDINAMARCA  ZIPAQUIRÁ    25899000          NaN  44245  FEMENINO

       GRUPO ETARIO  CANTIDAD
476968      ADULTOS         1
476969      ADULTOS         1
```

Figura 14. Listado los valores nulos del campo Armas Medios en el Archivo .csv
Fuente: elaboración propia

Se encuentran solo dos registros nulos en la variable de grupo etario.

En la figura 15, se realiza la conversión del Campo Fecha a un formato de día/Mes/Año en el reporte de violencia doméstica y se visualiza en el orden cronológico:

```
#Convertir Cadenas a Fecha y hora
datos['FECHA'] = pd.to_datetime(datos['FECHA'], format= "%d/%m/%Y", errors='coerce')
datos['FECHA'] = datos['FECHA'].fillna(0)
datos = datos[datos['FECHA'] != 0]
```

```
#imprimir resultados en orden cronologico la Fecha
datos.sort_values('FECHA')
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
0	ATLÁNTICO	BARRANQUILLA (CT)	8001000	ARMA BLANCA / CORTOPUNZANTE	2010-01-01 00:00:00	MASCULINO	ADULTOS	1
90	TOLIMA	IBAGUÉ (CT)	73001000	CONTUNDENTES	2010-01-01 00:00:00	FEMENINO	ADULTOS	3
89	SUCRE	SINCELEJO (CT)	70001000	CONTUNDENTES	2010-01-01 00:00:00	MASCULINO	ADULTOS	1
88	SUCRE	SINCELEJO (CT)	70001000	CONTUNDENTES	2010-01-01 00:00:00	FEMENINO	ADULTOS	5
87	SUCRE	SANTIAGO DE TOLÚ	70820000	CONTUNDENTES	2010-01-01 00:00:00	FEMENINO	ADULTOS	1
...	...	...	...	...	...	...	...	...
544916	MAGDALENA	San Zenón	47703000	SIN EMPLEO DE ARMAS	2023-03-31 00:00:00	FEMENINO	ADULTOS	1
542839	CÓRDOBA	Montería (CT)	23001000	SIN EMPLEO DE ARMAS	2023-03-31 00:00:00	FEMENINO	ADULTOS	1
544871	MAGDALENA	Ciénaga	47189000	SIN EMPLEO DE ARMAS	2023-03-31 00:00:00	FEMENINO	ADULTOS	1
542082	CALDAS	Manizales (CT)	17001000	SIN EMPLEO DE ARMAS	2023-03-31 00:00:00	FEMENINO	ADULTOS	2
548041	VICHADA	Santa Rosalía	99624000	SIN EMPLEO DE ARMAS	2023-03-31 00:00:00	FEMENINO	ADOLESCENTES	1

Figura 15. Cambio de formato de fecha y visualizar en orden cronológico el campo de fecha del Archivo .csv
Fuente: elaboración propia

En la figura 16, Se realiza el cambio del formato del campo Código Dane a entero y reemplazar el campo no reportado por el signo:

```
#Reemplaza el Codigo Dane a Entero
datos['CODIGO DANE'] = datos['CODIGO DANE'].replace({'NO REPORTADO':0,'-':0})
datos['CODIGO DANE'] = datos['CODIGO DANE'].astype(int)
datos = datos[datos['CODIGO DANE'] != 0]
```

Figura 16. Cambio de formato del código Dane y Reemplazar no reportado por el signo - del Archivo .csv
Fuente: elaboración propia

En la figura 17, Se muestra a información de los campos de las columnas del reporte de violencia domestica donde se evidencia código dane se encuentra en Tipo Entero:

```
#Información de tipos de dato del dataFrame
datos.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 548026 entries, 0 to 548041
Data columns (total 8 columns):
 #   Column         Non-Null Count   Dtype
---  ------         --------------   -----
 0   DEPARTAMENTO   548026 non-null  object
 1   MUNICIPIO      548026 non-null  object
 2   CODIGO DANE    548026 non-null  int32
 3   ARMAS MEDIOS   548026 non-null  object
 4   FECHA          548026 non-null  object
 5   GENERO         548023 non-null  object
 6   GRUPO ETARIO   546411 non-null  object
 7   CANTIDAD       548026 non-null  int64
dtypes: int32(1), int64(1), object(6)
memory usage: 35.5+ MB
```

Figura 17. Visualiza la información de los campos del Archivo .csv
Fuente: elaboración propia

En la figura 18, Se muestra en orden numérico el código de dane según el reporte de violencia doméstica:

```
#imprimir resultados en orden numerico el Codigo
datos.sort_values('CODIGO DANE')
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
138744	ANTIOQUIA	MEDELLÍN (CT)	5001000	CONTUNDENTES	2015-09-27 00:00:00	FEMENINO	MENORES	1
450805	ANTIOQUIA	MEDELLÍN (CT)	5001000	SIN EMPLEO DE ARMAS	2020-12-09 00:00:00	MASCULINO	ADULTOS	1
450804	ANTIOQUIA	MEDELLÍN (CT)	5001000	SIN EMPLEO DE ARMAS	2020-12-09 00:00:00	FEMENINO	ADULTOS	1
450803	ANTIOQUIA	MEDELLÍN (CT)	5001000	SIN EMPLEO DE ARMAS	2020-12-09 00:00:00	FEMENINO	ADULTOS	1
450802	ANTIOQUIA	MEDELLÍN (CT)	5001000	SIN EMPLEO DE ARMAS	2020-12-09 00:00:00	FEMENINO	ADULTOS	1
...	...	...	...	...	...	...	...	...
167190	VICHADA	CUMARIBO	99773000	CONTUNDENTES	2016-04-25 00:00:00	FEMENINO	ADULTOS	1
372423	VICHADA	CUMARIBO	99773000	ARMA BLANCA / CORTOPUNZANTE	2020-02-22 00:00:00	FEMENINO	ADULTOS	1
312147	VICHADA	CUMARIBO	99773000	NO REPORTADO	2019-04-12 00:00:00	FEMENINO	ADULTOS	1
534775	VICHADA	CUMARIBO	99773000	SIN EMPLEO DE ARMAS	2022-09-01 00:00:00	MASCULINO	ADULTOS	1
231915	VICHADA	CUMARIBO	99773000	SIN EMPLEO DE ARMAS	2017-08-10 00:00:00	MASCULINO	ADULTOS	1

548026 rows × 8 columns

Figura 18. Listado en orden numérico del Código Dane del Archivo .csv
Fuente: elaboración propia

En la figura 19. Se realiza la búsqueda de valores no reportados, los valores que contengan signo "- "son reemplazados por No reportado y se imprime la lista de géneros que se tiene en el reporte de violencia doméstica:

```
#rellenar los valores nulos con y reemplazarlos
datos['GENERO'] = datos['GENERO'].fillna('NO REPORTADO')
datos['GENERO'] = datos['GENERO'].replace({'-':'NO REPORTADO'})
datos['GENERO'].unique()
```

```
array(['MASCULINO', 'FEMENINO', 'NO REPORTADO'], dtype=object)
```

Figura 19. Búsqueda de valores no reportado, reemplazar e imprimir la lista de genero del campo Genero del Archivo .csv
Fuente: elaboración propia

En la figura 20, Se visualiza el listado en el orden alfabético del campo Genero en el reporte de Violencia Domestica:

```
#imprimir resultados en orden alfabetico
datos.sort_values('GENERO')
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
274013	CALDAS	ARANZAZU	17050000	ARMA BLANCA / CORTOPUNZANTE	2018-06-10 00:00:00	FEMENINO	ADULTOS	1
339082	PUTUMAYO	PUERTO ASÍS	86568000	NO REPORTADO	2019-10-06 00:00:00	FEMENINO	ADULTOS	2
339081	NORTE DE SANTANDER	OCAÑA	54498000	NO REPORTADO	2019-10-06 00:00:00	FEMENINO	ADULTOS	2
339079	NORTE DE SANTANDER	CÚCUTA (CT)	54001000	NO REPORTADO	2019-10-06 00:00:00	FEMENINO	ADULTOS	1
339078	NORTE DE SANTANDER	CÚCUTA (CT)	54001000	NO REPORTADO	2019-10-06 00:00:00	FEMENINO	ADOLESCENTES	1
...	...	...	...	...	...	...	...	...
517913	CUNDINAMARCA	SIBATÉ	25740000	SIN EMPLEO DE ARMAS	2022-05-05 00:00:00	NO REPORTADO	NO REPORTADO	1
524819	NARIÑO	SAN ANDRES DE TUMACO	52835000	SIN EMPLEO DE ARMAS	2022-06-06 00:00:00	NO REPORTADO	NO REPORTADO	1
502841	ANTIOQUIA	MEDELLÍN (CT)	5001000	SIN EMPLEO DE ARMAS	2022-12-07 00:00:00	NO REPORTADO	NO REPORTADO	1
514052	CUNDINAMARCA	BOGOTÁ D.C. (CT)	11001000	SIN EMPLEO DE ARMAS	2022-11-28 00:00:00	NO REPORTADO	NO REPORTADO	1
515403	CUNDINAMARCA	FUNZA	25286000	SIN EMPLEO DE ARMAS	2022-07-03 00:00:00	NO REPORTADO	NO REPORTADO	1

548026 rows × 8 columns

Figura 20. Listado en orden alfabético del Campo Género del Archivo .csv
Fuente: elaboración propia

En la figura 21, Se realiza la búsqueda de los valores no reportado del campo Armas medios y los que contengan signo – son reemplazados por No reportado y visualizar la lista de Armas Medio que contiene el campo en el reporte de violencia doméstica.

```
##rellenar los valores nulos con y reemplazarlos
datos['ARMAS MEDIOS'] = datos['ARMAS MEDIOS'].fillna('NO REPORTADO')
datos['ARMAS MEDIOS'] = datos['ARMAS MEDIOS'].replace({'-':'NO REPORTADO','NO REPORTADO':'NO REPORTADO'})
datos['ARMAS MEDIOS'].unique()
```

```
array(['ARMA BLANCA / CORTOPUNZANTE', 'ARMA DE FUEGO', 'CONTUNDENTES',
       'CORTANTES', 'CORTOPUNZANTES', 'NO REPORTADO', 'PUNZANTES',
       'SIN EMPLEO DE ARMAS', 'ESCOPOLAMINA'], dtype=object)
```

Figura 21. Búsqueda de valores no reportado, reemplazar e imprimir la lista de genero del campo Armas Medios del Archivo .csv
Fuente: elaboración propia

En la figura 22. Se visualiza el listado en el orden alfabético del campo armas medios en el reporte de Violencia Domestica:

```
#imprimir resultados en orden alfabetico
datos.sort_values('ARMAS MEDIOS')
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
0	ATLÁNTICO	BARRANQUILLA (CT)	8001000	ARMA BLANCA / CORTOPUNZANTE	2010-01-01 00:00:00	MASCULINO	ADULTOS	1
198782	META	VILLAVICENCIO (CT)	50001000	ARMA BLANCA / CORTOPUNZANTE	2016-12-20 00:00:00	FEMENINO	ADULTOS	1
198781	CUNDINAMARCA	BOGOTÁ D.C. (CT)	11001000	ARMA BLANCA / CORTOPUNZANTE	2016-12-20 00:00:00	MASCULINO	ADULTOS	1
198780	CAUCA	SANTANDER DE QUILICHAO	19698000	ARMA BLANCA / CORTOPUNZANTE	2016-12-20 00:00:00	FEMENINO	ADULTOS	2
198779	CASANARE	YOPAL (CT)	85001000	ARMA BLANCA / CORTOPUNZANTE	2016-12-20 00:00:00	MASCULINO	ADULTOS	1
...	...	...	...	...	...	...	...	...
352347	CUNDINAMARCA	BOGOTÁ D.C. (CT)	11001000	SIN EMPLEO DE ARMAS	2020-01-01 00:00:00	MASCULINO	ADULTOS	1
352348	CUNDINAMARCA	BOGOTÁ D.C. (CT)	11001000	SIN EMPLEO DE ARMAS	2020-01-01 00:00:00	MASCULINO	ADULTOS	1
126565	ANTIOQUIA	ITAGUI	5360000	SIN EMPLEO DE ARMAS	2015-06-25 00:00:00	FEMENINO	ADULTOS	2
126571	ANTIOQUIA	TARSO	5792000	SIN EMPLEO DE ARMAS	2015-06-25 00:00:00	FEMENINO	ADULTOS	1
548041	VICHADA	Santa Rosalía	99624000	SIN EMPLEO DE ARMAS	2023-03-31 00:00:00	FEMENINO	ADOLESCENTES	1

Figura 22. Listado en orden alfabético del Campo armas medios del Archivo .csv
Fuente: elaboración propia

En la figura 23, Se realiza la búsqueda de valores nulos que son reemplazados por No Reportado en el campo de Grupo Etario en el reporte de violencia doméstica:

```
#sustituir los valores nulos por otro valor
datos['GRUPO ETARIO'] = datos['GRUPO ETARIO'].fillna('NO REPORTADO')
```

Figura 23. Reemplazo de valores nulos en el campo grupo Etario del Archivo .csv
Fuente: elaboración propia

En la figura 24, Mediante el grafico de histograma se visualizar los valores Adultos, géneros, menores y no reportado de manera correcta sin ningún nulo en el campo de grupo etario del reporte de violencia doméstica.

```
# Visualizar un histograma para validar poner el valor mas repetido
hist = datos['GRUPO ETARIO'].hist()
```

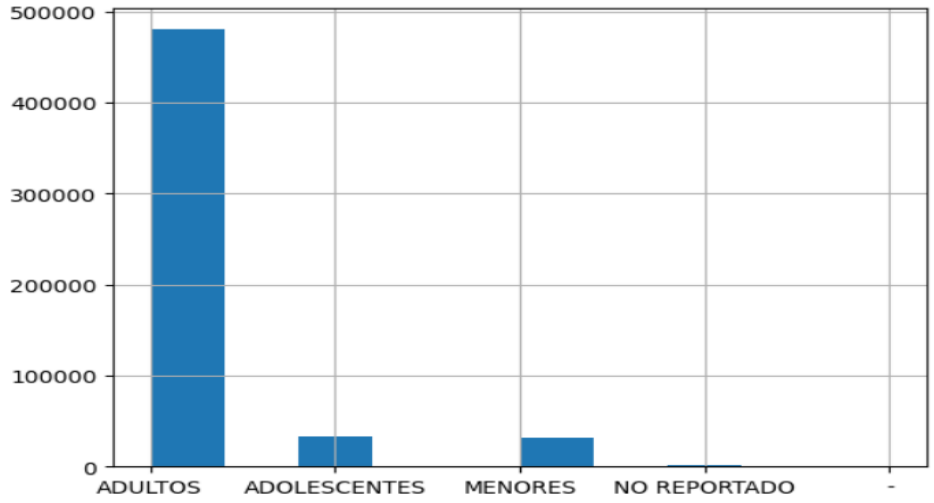


Figura 24. Listado general de los datos del Archivo .csv
Fuente: elaboración propia

En la figura 25, Se visualiza el listado en el orden alfabético del campo grupo etario en el reporte de Violencia Domestica:

```
#imprimir resultados en orden alfabetico
datos.sort_values('GRUPO ETARIO')
```

	DEPARTAMENTO	MUNICIPIO	CODIGO DANE	ARMAS MEDIOS	FECHA	GENERO	GRUPO ETARIO	CANTIDAD
543279	CUNDINAMARCA	Bogotá D.C. (CT)	11001000	SIN EMPLEO DE ARMAS	2023-02-28 00:00:00	NO REPORTA	-	1
543365	CUNDINAMARCA	Bogotá D.C. (CT)	11001000	SIN EMPLEO DE ARMAS	2023-03-14 00:00:00	NO REPORTA	-	1
543167	CUNDINAMARCA	Bogotá D.C. (CT)	11001000	SIN EMPLEO DE ARMAS	2023-02-10 00:00:00	NO REPORTA	-	1
539911	ANTIOQUIA	Copacabana	5212000	SIN EMPLEO DE ARMAS	2023-03-08 00:00:00	NO REPORTA	-	1
543372	CUNDINAMARCA	Bogotá D.C. (CT)	11001000	SIN EMPLEO DE ARMAS	2023-03-15 00:00:00	NO REPORTA	-	1
...	...	...	...	...	...	...	...	...
502713	ANTIOQUIA	MEDELLÍN (CT)	5001000	SIN EMPLEO DE ARMAS	2022-10-30 00:00:00	NO REPORTADO	NO REPORTADO	1
520179	GUAJIRA	MAICAO	44430000	SIN EMPLEO DE ARMAS	2022-02-23 00:00:00	NO REPORTADO	NO REPORTADO	1
502704	ANTIOQUIA	MEDELLÍN (CT)	5001000	SIN EMPLEO DE ARMAS	2022-10-28 00:00:00	NO REPORTADO	NO REPORTADO	1
513946	CUNDINAMARCA	BOGOTÁ D.C. (CT)	11001000	SIN EMPLEO DE ARMAS	2022-11-11 00:00:00	NO REPORTADO	NO REPORTADO	1
520178	GUAJIRA	MAICAO	44430000	SIN EMPLEO DE ARMAS	2022-02-14 00:00:00	NO REPORTADO	NO REPORTADO	1

548026 rows × 8 columns

Figura 25. Listado en orden alfabético del campo grupo etario del Archivo .csv
Fuente: elaboración propia

En la figura 26, Se muestra en la figura el listado de los campos que ya no contiene valores nulos en el reporte de violencia doméstica.

```
#Revisar nuevamente si presenta valores Nulos
datos.isnull().sum()
```

```
DEPARTAMENTO    0
MUNICIPIO       0
CODIGO DANE     0
ARMAS MEDIOS    0
FECHA           0
GENERO          0
GRUPO ETARIO    0
CANTIDAD        0
dtype: int64
```

Figura 26. Listado general de los datos del Archivo .csv
Fuente: elaboración propia

En la figura 27, Finalmente se exporta el archivo que paso por las técnicas de limpieza y extracción de los datos para luego ser entregado a la herramienta de Power bi para su visualización.

```
datos.to_csv("ReporteViolenciadomesticaPNCLimpia.csv")
```

Figura 27. Generar Archivo .csv
Fuente: elaboración propia

Visualización y análisis de los datos mediante Power Bi

El siguiente tablero de control desarrollado por la aplicación de Power Bi en donde se va a identificar los patrones y tendencias de la violencia doméstica en Colombia con los datos públicos de la policía nacional de Colombia.

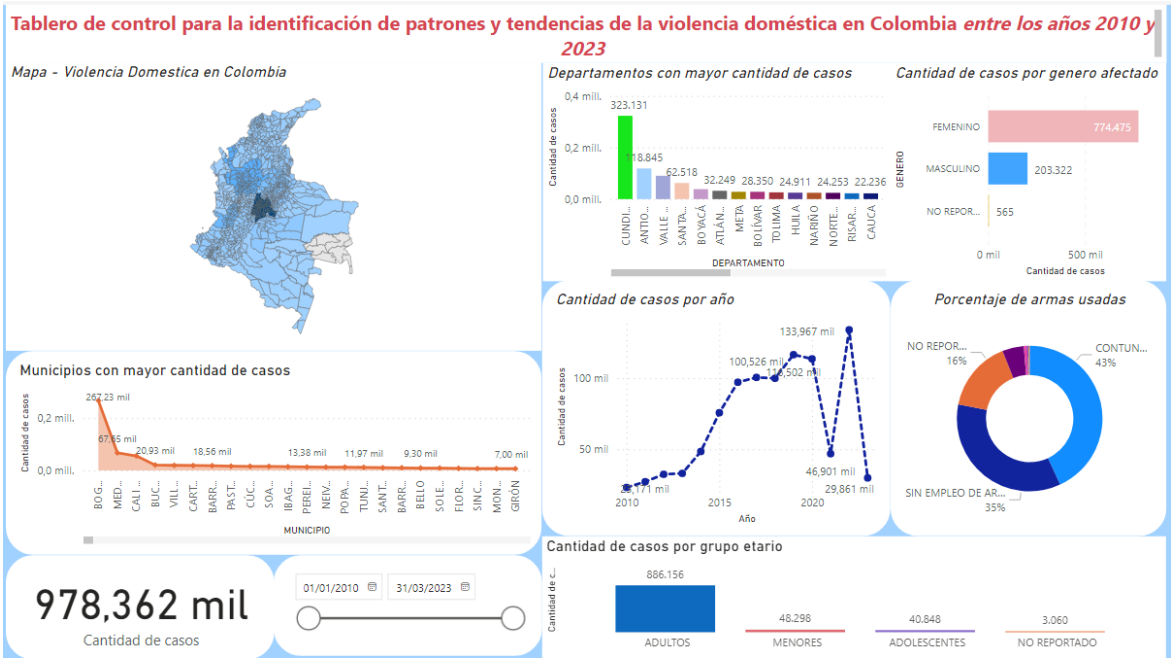


Figura 28. Tablero de Control mediante Power Bi.
Fuente: elaboración propia

En la figura 29, se genera un mapa de Colombia en donde se visualiza los departamentos con casos de violencia doméstica y al sobreponer en un departamento nos indica la cantidad de casos. Al ser mayor el color también nos indica que es mayor el índice de violencia en esa zona.

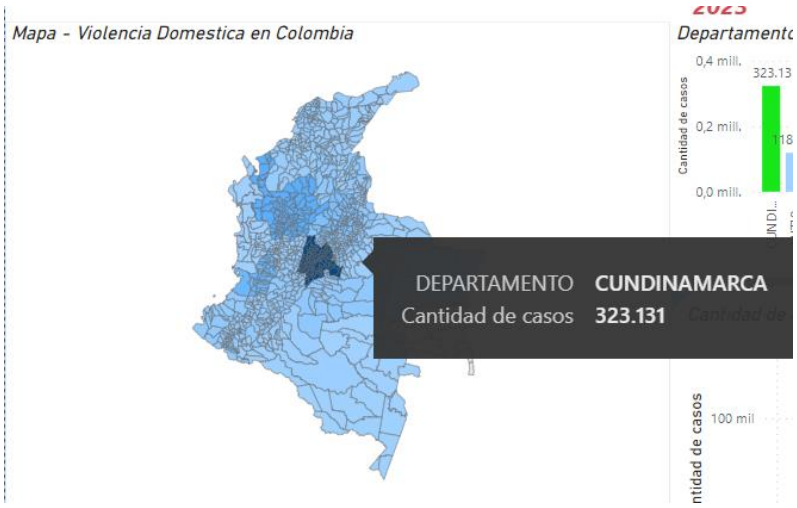


Figura 29. Mapa de Colombia visualiza la cantidad de casos por departamento.
Fuente: elaboración propia

Esta variable resulta de gran relevancia en este caso debido a que la violencia en Colombia es generada desde varios grupos subversivos que surgen desde diferentes

regiones del país que pueden aportar de manera significativa a la violencia intrafamiliar, adicional, con esta información sería fácil accionar una estrategia por regiones direccionadas a diferentes problemáticas que se presenten, ya sea falta de educación, recursos económicos o dominancia de grupos bélicos que coaccionan a la comunidad para mantener un yugo y una sumisión que se puede ver reflejada de manera negativa en los hogares.

"Entre 1974 y 1990 la historia de Colombia se encuentra marcada por varias reformas constitucionales que, a pesar de presentar algunos avances, mantendría serias deficiencias sociales de carácter estructural. Entre ellas destaca la promulgación de una reforma agraria con el objetivo de permitir el acceso a las tierras. En 1978 la represión política y armada ejercida bajo la legislatura del presidente Julio César Turbay Ayala (1978-1982) del Partido Liberal fue conocida por la promulgación del controvertido Estatuto de Seguridad bajo el objetivo de apaciguar las revueltas sociales que desencadenó numerosas violaciones a los derechos humanos. En paralelo, a lo largo de la década de 1970 se realizaron cambios drásticos en política económica: se redujo la intervención estatal dando lugar a los procesos de liberalización económica. El resultado fue el deterioro de la estructura social, el aumento de las tasas de interés, aumento de la especulación y la concentración empresarial" (Cidob.org, 2014).

En la siguiente figura 30, nos visualiza los departamentos con mayor índice de casos de violencia doméstica en el Colombia.

Esta visualización permite que las organizaciones con esta información puedan tomar acciones e intervenir los departamentos que mayor número de casos de violencia tenga y de esta manera poder asignar recursos

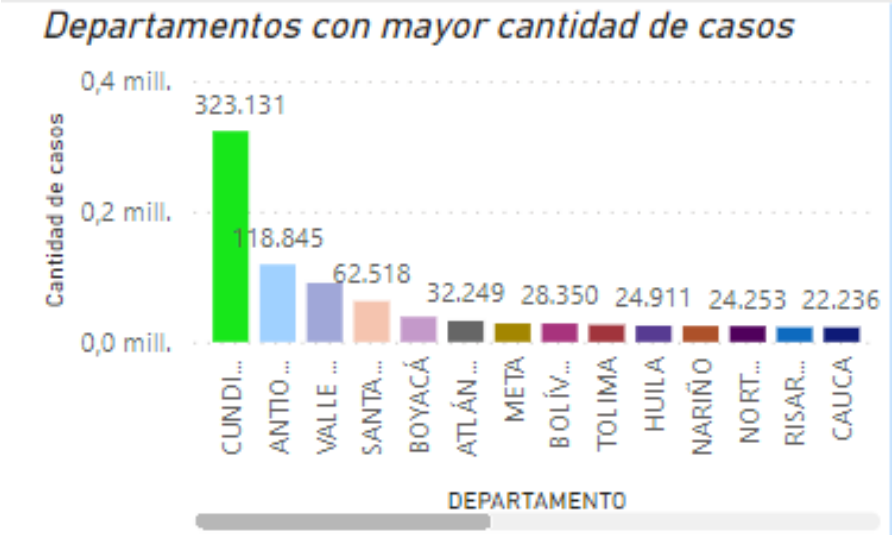


Figura 30. Departamentos con mayor cantidad de casos.
Fuente: elaboración propia

En la figura 31, se visualiza cantidad de casos por genero afectado por la violencia doméstica en Colombia.

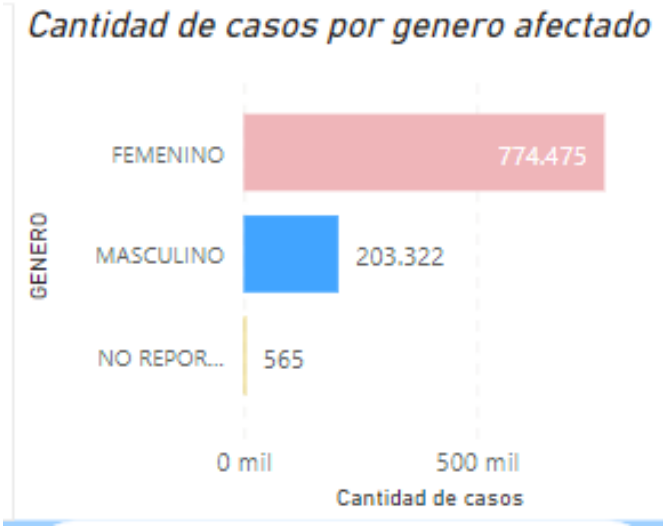


Figura 31. Cantidad de Casos por genero afectado
Fuente: elaboración propia

La variable de género resulta ser la más importante para este estudio y esta gráfica nos muestra de manera clara el afán de implementar estrategias fuertes para establecer seguridad en la mujer, educación y protección; así mismo, por regiones se podría determinar si esto

depende de razones culturales, sociales o psicológicas desde otras generaciones arraigadas en la violencia a la mujer.

En la figura 32, se visualiza las armas más usadas para generar los indicadores de violencia doméstica en Colombia.

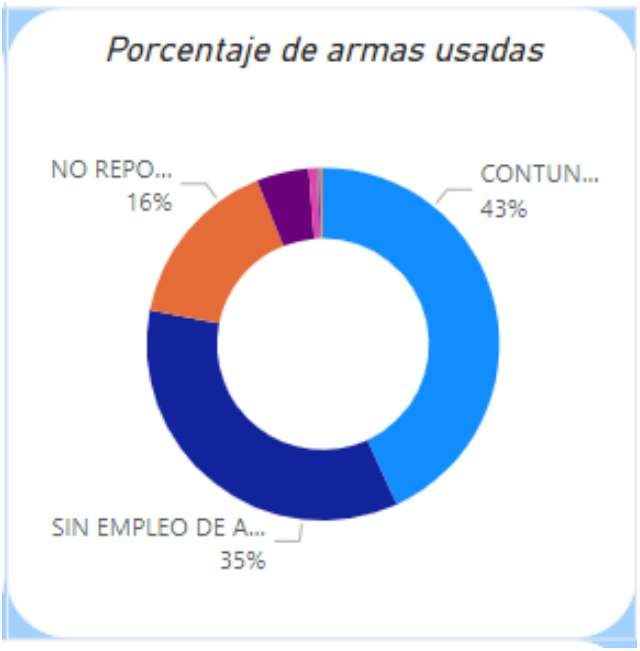


Figura 32. Porcentaje de armas usadas
Fuente: elaboración propia

En este caso el tipo de armas como variable de análisis puede ayudarnos a definir el nivel de la violencia generada, si se genera desde grupos delincuenciales comunes o grupos bélicos establecidos con entrenamiento e inventario en armas y relacionados.

En la figura 33, se visualiza la cantidad de casos por año que se viene registrando por violencia doméstica en Colombia.



Figura 33.Cantidad de casos por Año
Fuente: elaboración propia

En la figura 34, se visualiza los municipios o capitales de Colombia donde se tiene un mayor índice de violencia doméstica. Al pasar el mouse por la gráfica nos indica la cantidad que casos exactos.



Figura 34.Municipios con mayor cantidad de casos
Fuente: elaboración propia

En la figura 35, se visualiza el grupo etario donde los adultos se tiene un mayor índice de violencia doméstica.

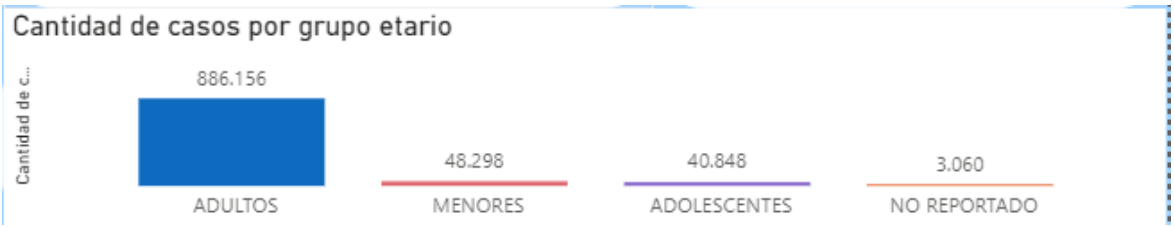


Figura 35. Cantidad de Casos por grupo etario
Fuente: elaboración propia

3. Resultados

Diseño del Dashboard con la aplicación de Power BI

En la figura 36, Se construye tablero de control para se va a identificar los patrones y tendencias de la violencia doméstica en Colombia.

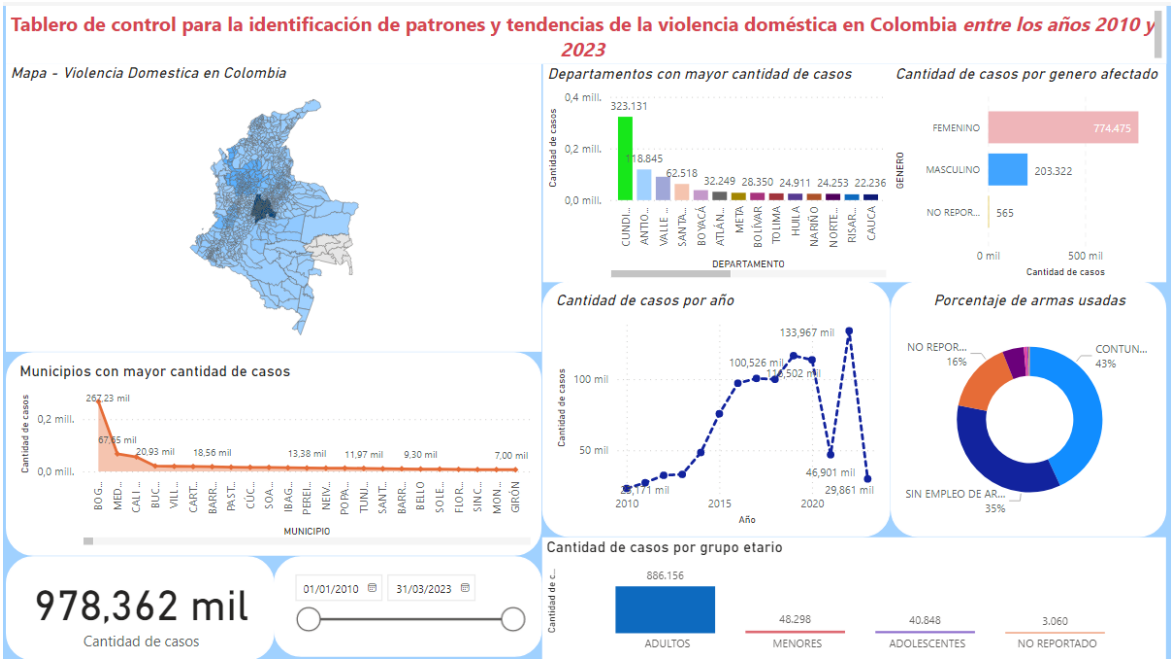


Figura 36.Departamentos con mayor cantidad de casos.
Fuente: elaboración propia

En la figura 37, se puede observar en los 32 departamentos de Colombia desde 2010 al marzo 2023, la cantidad de casos varió de 485 a 323.131 mil reportados de violencia doméstica en Colombia. El departamento de Cundinamarca reporta la mayor cantidad de casos, con un total de 323.131 mil, lo que corresponde al 33% y supera a Vichada, el departamento con menos casos reportados, que registra 485.

En este caso también debemos analizar otras subvariables como la cantidad per cápita de cada departamento y encontrar la proporción correspondiente en habitantes para poder definir que realmente Cundinamarca pueda ser más violenta en el ámbito intrafamiliar; evidentemente si

será mayor en cantidad dado que corresponde a los departamentos que son los de mayor población.

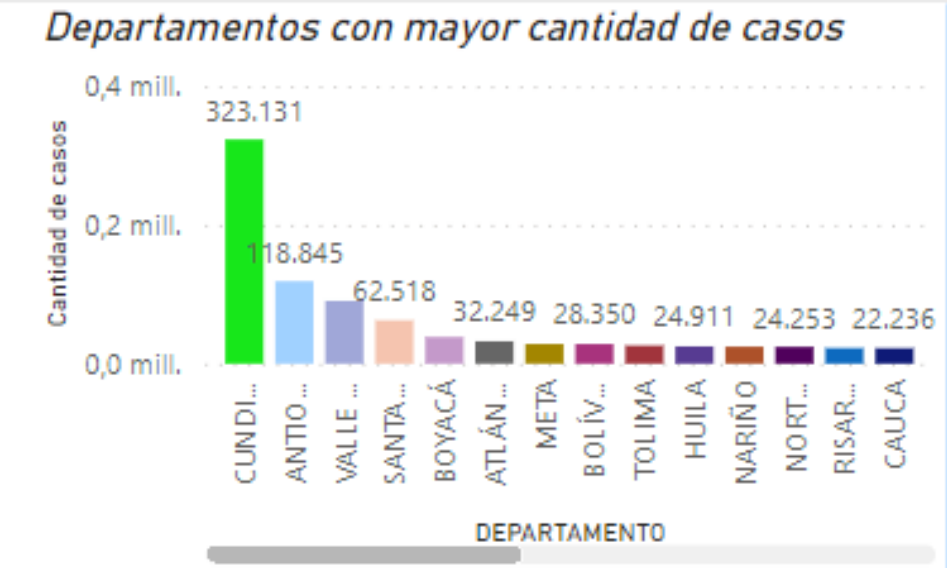


Figura 37. Departamentos con mayor cantidad de casos.
Fuente: elaboración propia

En la figura 38, se observa que el género femenino lidera un número de casos, con un total de 774.475, seguido por el género masculino con 203.322, mientras que el grupo "no reportado" que cuenta con 565 casos. El género femenino representa con el 79,16% de los casos con mayor violencia doméstica en Colombia durante el periodo del 2010 al marzo 2023.

Dejando como referencia una clara necesidad de una estrategia que sea de apoyo a la mujer en todos los casos posibles sin discriminar raza, creencia, etc., prácticamente la relación resulta ser que por cada hombre violentado hay 3 mujeres violentadas en Colombia y por lo menos por los últimos 12 años, según la estadística y la información analizada.

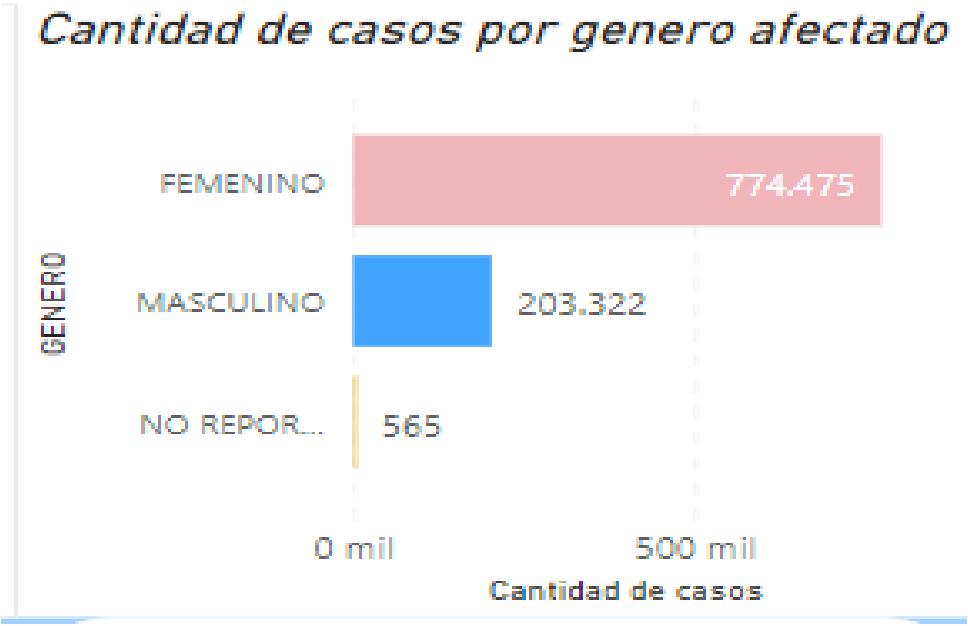


Figura 38.Cantidad de Casos por genero afectado
Fuente: elaboración propia

En la figura 39, se observa el arma medios más utilizada para cometer los actos de violencia domestica es las armas contundentes que corresponde al 43%.

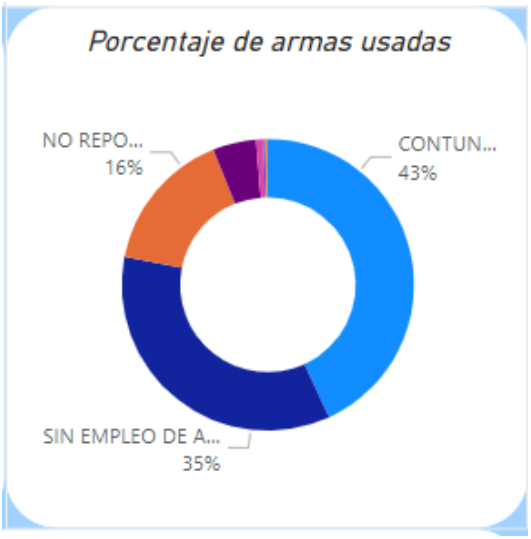


Figura 39.Porcentaje de armas usadas
Fuente: elaboración propia

De esta representación gráfica podemos deducir claramente el tipo de arma utilizado y de allí poder entender si la legalización de las mismas afecta o no en ciertas regiones del país más que en otras, entender, quiénes están teniendo acceso a ellas y de qué forma, cómo se mueve el tráfico de armas y cómo los hombres principalmente pueden acceder a ellas, esto debido a que las victimas principalmente por ejemplo por arma contundente son los hombres, en cambio, las mujeres son atacadas sin ellas o no reportan principalmente, cuánto se pueden ver afectados los niños o los adultos.

En la figura 40, se observa que la cantidad de casos de violencia domestica dio un inicio con tendencia al alza partir del año 2014 con 48.440, siendo el año 2022 con 133.967 con el índice más alto que corresponde al 13,69 % en los siguientes 8 años.



Figura 40.Cantidad de casos por Año
Fuente: elaboración propia

Evidenciamos un comportamiento violento también depende de las situaciones que a nivel nacional o mundial estén sucediente, en el caso de la pandemia encontramos un comportamiento demasiado atípico en el que la violencia intrafamiliar aumentó casi el doble y esto dependió de muchos factores sociales, económicos, emocionales y psicológicos.

En el año 2021 también vemos una disminución significativa y podría deberse a que la comunidad tomó conciencia del valor de la vida, probablemente hubo mayor asistencia

psicológica y médica o es también es probable revisar si hubo crisis jurídica y por ello no fue posible realizar denuncias con prontitud.

En la figura 41, se observa en la capital de Colombia Bogotá D.C. corresponde el 27,31% de los demás departamentos con más casos reportados de violencia doméstica. En los 1.022 municipios de Colombia, la cantidad de casos reportados de violencia domestica varió de 2 a 267.228mil.



Figura 41.Municipios con mayor cantidad de casos
Fuente: elaboración propia

En los casos en que el arma es contundente sucede un comportamiento muy curioso en donde nos muestra que las cifras son mayores en los departamentos de Antioquia y Valle del Cauca que en el departamento de Cundinamarca, aunque este último tenga la mayor parte de las cifras, no es el que mayor número de víctimas que aporta en este aspecto, dándonos muestra de comportamientos particulares que se presentan en estas regiones y que pueden ser erradicados de raíz con programas fuertes en toda la comunidad paisa y Valluna.

En la figura 42, se observa que los adultos corresponden al 90,58% con los 886.156 mil casos reportados violencia doméstica en Colombia entre los años 2010 y marzo del 2023. También se tiene que los 4 grupos etarios la cantidad de casos vario de 3060 a 886.156.

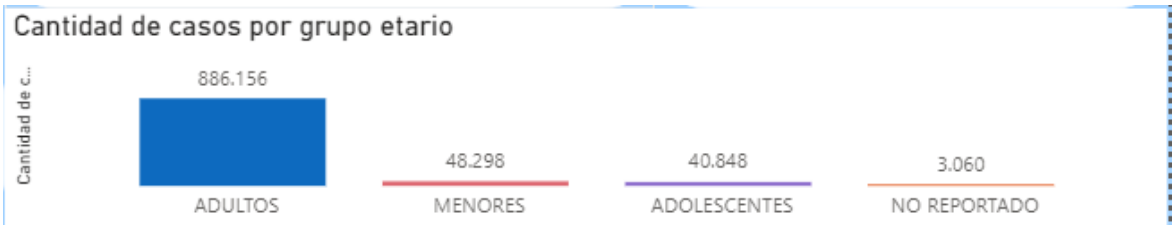


Figura 42. Cantidad de Casos por grupo etario
Fuente: elaboración propia

Es así como a través de este cuadro de control podemos evidenciar datos que son supuestos y aquí se materializan al saber que son reales y que se pueden cruzar entre ellas mismas, por ejemplo, el saber que una mujer es más violentada, por qué razón o que afectan esos resultados, es de vital importancia para entidades, empresas y personas que puedan aportar y crear estrategias de solución a esta problemática tan importante en la comunidad colombiana y para el Estado colombiano.

Una vez finalizada la creación del tablero de visualización de datos sobre las estadísticas de violencia doméstica, se llevó a cabo un proceso de validación por parte de un experto temático en el campo. El objetivo de esta etapa fue evaluar la efectividad y la pertinencia del tablero en cuanto a su capacidad para proporcionar una visualización clara y comprensible de los datos relacionados con la violencia doméstica. El experto temático seleccionado cuenta con experiencia y se encuentra familiarizado con los indicadores y las métricas utilizadas en el campo. Se le proporcionó acceso al tablero y se le solicitó que evaluara diferentes aspectos, como la presentación de los datos, la facilidad de navegación, la claridad de los gráficos y la relevancia de las visualizaciones para comprender el problema de la violencia doméstica, y en términos generales indica que este tipo de tableros son de importancia dentro de su gremio a razón que les permite realizar consultas de información real sobre la violencia doméstica (ver anexo 1).

4. Discusión

La presente investigación ha proporcionado resultados significativos sobre la incidencia de la violencia doméstica en Colombia. Las conclusiones obtenidas a partir del análisis de los datos muestran la necesidad de tomar medidas efectivas para abordar este grave problema social. A continuación, se discutirán las implicaciones de los resultados y se destacarán los aportes más importantes de la investigación.

En primer lugar, se destaca que el género femenino presenta la mayor cantidad de casos de violencia doméstica, representando el 79,16% de los casos reportados. Esto resalta la vulnerabilidad de las mujeres ante este tipo de violencia y la necesidad de implementar estrategias específicas para su protección y apoyo. Estos resultados coinciden con investigaciones previas que demuestran que las mujeres son las principales víctimas de la violencia doméstica.

Por otro lado, se observa que el departamento de Cundinamarca cuenta con la mayor cantidad de casos reportados, representando el 33% del total. Esto indica la necesidad de enfocar esfuerzos y recursos en esta región para abordar la violencia doméstica de manera efectiva. Además, se resalta la importancia de prestar atención a los departamentos con menor cantidad de casos reportados, como Vaupés, para garantizar que también se implementen medidas de prevención y apoyo en estas áreas.

Otro hallazgo relevante es el hecho de que el arma más utilizada para cometer actos de violencia doméstica es el arma contundente, que representa el 43% de los casos. Estos resultados señalan la importancia de abordar el acceso y uso de armas en el ámbito doméstico, así como implementar medidas de prevención y concienciación sobre los riesgos asociados con su uso en situaciones de violencia familiar.

La tendencia al alza en la cantidad de casos de violencia doméstica desde 2014, con un aumento del 38,35% en los siguientes 9 años, es un dato alarmante que resalta la necesidad de actuar de manera urgente. Estos resultados indican que los esfuerzos actuales no están siendo suficientes para frenar este problema, por lo que se requiere una mayor atención y recursos para abordar eficazmente la violencia doméstica en el país.

En cuanto a las recomendaciones, se destacan varias acciones clave. En primer lugar, es necesario fortalecer los esfuerzos de sensibilización y educación en torno a la violencia doméstica, tanto para las posibles víctimas como para la sociedad en general. Esto implica realizar campañas de concienciación y talleres educativos, así como programas de capacitación para profesionales que trabajan en este ámbito.

Además, se resalta la importancia de fortalecer los servicios de apoyo a las víctimas de violencia doméstica. Esto implica garantizar que existan recursos adecuados, como líneas de ayuda telefónica, refugios seguros y asesoramiento psicológico y legal, para brindar un apoyo integral a las víctimas.

La coordinación entre instituciones y agencias relevantes es fundamental para garantizar una respuesta efectiva a la violencia doméstica. Esto implica establecer protocolos claros de colaboración, compartir información de manera segura y coordinar esfuerzos en la prevención, protección y persecución de los casos.

Es necesario invertir en programas de sensibilización y capacitación para profesionales de la salud, la educación y el sistema de justicia, para que estén preparados para identificar y abordar adecuadamente los casos de violencia doméstica.

5. Conclusiones y recomendaciones

5.1 Conclusiones

La utilización de Python y Power BI en la investigación sobre la violencia doméstica ha brindado varios beneficios y aportes significativos. Estas herramientas han permitido un enfoque más eficiente y preciso en el tratamiento y análisis de los datos, así como una visualización clara y comprensible de los resultados obtenidos. Algunos puntos para rescatar frente al uso de estas herramientas son:

Automatización y eficiencia del análisis: Python es un lenguaje de programación poderoso y versátil ha permitido procesar y analizar grandes volúmenes de datos de manera eficiente. Su amplia gama de bibliotecas y herramientas específicas para el análisis de datos te han facilitado la manipulación y limpieza de los datos, así como la realización de cálculos y estadísticas relevantes para la investigación. Esto ha agilizado el proceso de análisis y ha permitido obtener resultados más rápidos y confiables.

Flexibilidad en la visualización de resultados: Power BI es una herramienta de visualización de datos altamente intuitiva y poderosa. Al utilizarla en la investigación, se pudo crear gráficos, tablas y paneles interactivos que facilitan la comprensión y comunicación de los hallazgos. La capacidad de personalización de Power BI permitió presentar los datos de manera clara y atractiva, lo cual es fundamental para transmitir eficazmente los resultados a diferentes audiencias, como investigadores, responsables de políticas públicas o el público en general.

Identificación de patrones y tendencias: El análisis de datos realizado con Python y la visualización en Power BI han posibilitado identificar patrones y tendencias en la violencia doméstica. Estas herramientas han permitido explorar y analizar datos de diferentes variables relevantes, como el género, la edad, el tipo de violencia, entre otros. Al identificar patrones y tendencias, se pudo obtener información valiosa para comprender mejor los factores subyacentes, lo cual contribuye a la generación de conocimiento y a la toma de decisiones informadas para abordar este grave problema social.

Reproducibilidad y transparencia: El uso de Python para el tratamiento de los datos ofrece ventajas en términos de reproducibilidad y transparencia. Al utilizar código abierto y compartir los scripts y métodos utilizados, otros investigadores y profesionales pueden replicar los análisis y verificar los resultados. Esto fomenta la transparencia en la investigación y permite la colaboración y mejora continua en el campo de estudio de la violencia doméstica.

Basándonos en los datos proporcionados sobre la problemática de la violencia doméstica en Colombia, se pueden obtener las siguientes conclusiones y orientaciones importantes para brindar soluciones y generar toma de decisiones:

El género femenino representa la mayor proporción de casos de violencia doméstica en Colombia. Esta conclusión resalta la necesidad de focalizar los esfuerzos en la protección y apoyo a las mujeres que sufren violencia en el ámbito doméstico. Las políticas y programas deben considerar la perspectiva de género y ofrecer servicios específicos para atender las necesidades de las mujeres afectadas.

El departamento de Cundinamarca concentra una alta cantidad de casos reportados de violencia doméstica. Esto indica la necesidad de priorizar medidas preventivas en esta región, como campañas de concienciación, educación y fortalecimiento de los sistemas de apoyo a las víctimas. Además, es fundamental establecer alianzas con las autoridades locales y la sociedad civil para abordar este problema de manera integral.

La tendencia creciente de casos de violencia doméstica en los últimos años indica que se requiere una respuesta sostenida y continua para abordar este problema. Es necesario fortalecer las iniciativas existentes y generar nuevas estrategias que involucren a diversos actores, como instituciones gubernamentales, organizaciones no gubernamentales y la sociedad civil en general.

La concentración de casos en Bogotá D.C. resalta la importancia de implementar políticas y programas específicos en la capital del país. Estos deben abordar las causas subyacentes de la violencia doméstica y fortalecer los servicios de apoyo a las víctimas en esta área. También es fundamental fortalecer la coordinación entre las diferentes

entidades y organismos encargados de la protección de los derechos de las mujeres en la ciudad.

Los adultos son el grupo de edad más afectado por la violencia doméstica en Colombia. Esto destaca la importancia de diseñar políticas y programas que aborden las necesidades específicas de este grupo, como la atención psicosocial, el acceso a refugios seguros y la capacitación laboral para ayudar a las víctimas a reconstruir sus vidas.

La propuesta generada en el presente documento permite evidenciar y visualizar las variables críticas de violencia doméstica en Colombia, dando una herramienta para la toma de decisiones con el fin de que puedan desarrollar análisis más profundos. Esto soportado en los aportes del experto temático que fueron considerados para realizar ajustes y mejoras al tablero en base a sus recomendaciones. Esto incluyó modificaciones en la presentación gráfica, la incorporación de filtros y dimensiones adicionales, así como la inclusión de métricas clave que facilitaran el análisis y la comprensión de los datos.

5.2 Recomendaciones

A continuación, se exponen una serie de aspectos y/o elementos que se podrían realizar en un futuro para emprender proyectos similares o fortalecer el proceso realizado.

Es importante continuar con los esfuerzos de sensibilización y educación en torno a la violencia doméstica, tanto para las víctimas potenciales como para la sociedad en general. Esto puede incluir campañas de concienciación, talleres educativos en escuelas y comunidades, y programas de capacitación para profesionales que trabajan en el campo de la violencia doméstica.

Fortalecer los servicios de apoyo a las víctimas de violencia doméstica es fundamental. Esto implica garantizar que existan recursos suficientes, como líneas de ayuda telefónica, refugios seguros, asesoramiento psicológico y asistencia legal, para brindar apoyo integral a las víctimas.

Fomentar la coordinación entre instituciones y agencias relevantes, como la policía, el sistema de justicia, los servicios sociales y los grupos de apoyo, es fundamental para garantizar una respuesta efectiva a la violencia doméstica. Esto implica establecer protocolos claros de colaboración, compartir información de manera segura y coordinar esfuerzos en la prevención, protección y persecución de los casos.

Se debe realizar una mayor inversión en programas de sensibilización y capacitación para profesionales de la salud, la educación y el sistema de justicia, para que puedan identificar y abordar de manera adecuada los casos de violencia doméstica.

Es importante promover la participación de la sociedad civil y las organizaciones no gubernamentales en la prevención y atención de la violencia doméstica, creando redes de apoyo y promoviendo la denuncia de los casos.

Se debe impulsar la investigación y recopilación de datos actualizados sobre la violencia doméstica, para contar con información precisa que permita la toma de decisiones informadas y la evaluación de las políticas implementadas.

Bibliografía

Amazon Web Services,2023, ¿Qué es Python? https://aws.amazon.com/es/what-is/python/#:~:text=Python%20es%20un%20lenguaje%20de,ejecutar%20en%20muchas%20plataformas%20diferentes.

Castaño-Monsalve, C. E., Ochoa-Montoya, J. A., & Vélez Álvarez, C. A. (2019). Barreras para denunciar la violencia intrafamiliar en Colombia. Revista de Salud Pública, 21(3), 365-370.
CIDOB. (2019), Conflicto en Colombia: antecedentes históricos y actores,https://www.cidob.org/publicaciones/documentacion/dossiers/dossier_proceso_de_paz_en_colombia/dossier_proceso_de_paz_en_colombia/conflicto_en_colombia_antecedentes_historicos_y_actores#:~:text=Entre%20las%20guerrillas%20m%C3%A1s%20notorias,Abril%20(M%2D19).

Dare Planet Technology. (2022). Análisis de datos mediante la librería Pandas de Python, https://dareplanet.tech/insights/analisis-de-datos-libreria-pandas-python/
Dressel, J. y. (2018). ScienceAdvances. Obtenido de https://www.science.org/doi/10.1126/sciadv.aao5580

Gómez, J. (2020). Factores de riesgo asociados a la violencia doméstica en Colombia. Revista Colombiana de Psicología, 29(1), 45-55.

González Ortiz Jorge O. (2002), VIOLENCIA INTRAFAMILIAR: UNA FORMA DE RELACIÓN, UN ASUNTO DE DERECHOS HUMANOS, https://www.medicinalegal.gov.co/documents/20143/49484/Violencia+Intrafamiliar.pdf/e634f6aa-905e-4f87-4084-921a4917ea8b

Instituto Colombiano de Bienestar Familiar, “Proceso gestión de atención a peticiones, quejas, reclamos y sugerencias”, http://www.icbf.gov.co/portal/page/portal/IntranetICBF/macro_procesos/MP_apoyo/G_

Instituto Nacional de Medicina Legal y Ciencias Forenses. (2021). Forensis 2021: estadísticas de violencia intrafamiliar. https://www.medicinalegal.gov.co/documents/20143/878249/Forensis_2021.pdf

John Keane, Reflexiones sobre la violencia, traducción de Josefa Linares de la Puerta, Madrid, Alianza editorial, 2000, pp. 61-62

Martha Báez. (México D. F.: Trillas, 2011). Aspectos de la violencia intrafamiliar a puerta cerrada,143.

Microsoft,2023, ¿Que es Power BI?,https://powerbi.microsoft.com/es-es/what-is-power-bi/.

Mora-Rojas, E., González-Barrera, D., & Gutiérrez-Cantero, O. (2020). Aplicación de minería de datos para la identificación de patrones y tendencias en la violencia intrafamiliar en Colombia. En Congreso Internacional de Investigación Academia Journals (pp. 128-136).
Oviedo, C. A., & Zapata, L. G. (2021).

Análisis de la violencia doméstica en Colombia mediante técnicas de minería de datos. Ingeniería y Competitividad, 23(1), 75-89.

Personería de Bogotá, Informe sobre los derechos de las mujeres en Bogotá. Derechos humanos: defenderlos, protegerlos, promoverlos, responsabilidad de todos y todas (Bogotá: Personaría de Bogotá, 2003), 34.

Policía Nacional de Colombia. (2023). Violencia Intrafamiliar. https://www.policia.gov.co/delitos-de-impacto/violencia-intrafamiliar.
Rachel Faria, &. S. (2016). Scielo.org,co. Obtenido de https://doi-org.crai-ustadigital.usantotomas.edu.co/10.1590/2175-623653298

Ramírez, F. A. (03 de 2021). monografias. Obtenido de https://www.monografias.com/trabajos59/antecedentes-historicos-violencia-familiar/antecedentes-historicos-violencia-familiar

Ridgeway, G. (2018). Policing in the Era of Big Data. En G. Ridgeway, Annual Review of Criminology (págs. pp. 401-419).

Rodríguez, L. F., & Zuleta, J. C. (2020). La violencia doméstica en Colombia: Una revisión sistemática de la literatura. Revista Salud & Sociedad, 11(2), 173-187.

Roy Rosenzweig Center for History and New Media. (09 de 2016). Zotero. Obtenido de https://www.zotero.org

Sáenz, J. (09 de 09 de 2016). Wiki Sistemas Digitales. Obtenido de http://sistdig.wikidot.com servicioyatencion/G_atencionpqrs/Procedimientos/G1.PR1.MPA5.P2%20%20Gu%C3%ADa%20de%20Constataci%C3%B3n%20de%20Denuncias%20v1.pdf (acceso agosto 7, 2013).

Suárez, R. (2015). Energías renovables, Impacto y Efectividad. Revista Científica Ingeniería y Desarrollo en Energía, 48-53.

Thomson Reuters. (09 de 2016). EndNote. Obtenido de http://endnote.com/

World Health Organization (who). (2014) Global Status Report on Violence Prevention. who. Washington, ee. uu. Recuperado de:

https://oig.cepal.org/sites/default/files/informe_sobre_la_situacion_mundial_de_la_prevencion_de_la_violencia.pdf

Anexo 1

Bogotá, 25 de mayo de 2023.

Señores

Fundación Universitaria Compensar

Seminario de grado

Especialización Big Data

Grupo de trabajo de grado: Tablero de control para la identificación de patrones y tendencias de la violencia doméstica en Colombia entre los años 2010 y 2022.

ASUNTO : Evaluar la efectividad del tablero de control con la identificación de patrones y tendencias en la violencia doméstica en Colombia y su capacidad para generar información útil para el diseño de estrategias de prevención y atención.

Cordial saludo,

Se realiza revisión del tablero de control según se referencia y encontramos que las variables que nos aportan al análisis son los siguientes:

1. Ubicación geográfica.
2. Clasificación por género.
3. Grupo etario.
4. Comportamiento histórico.
5. Arma utilizada.

Los cuales son de carácter relevante en la definición de comportamientos y patrones que nos aporten en el análisis, evaluación y generación de resultados que nos permitan construir estrategias focalizadas a la problemática en particular, por ejemplo:

1. Revisión de violencia en la mujer: Se evidencia que la cantidad de casos triplica en los casos que se reportan a hombres, es decir, esto ya nos da un claro argumento para seguir realizando actividades educativas en la formación de valores, respeto y apoyo a la mujer en todas las estancias y edades, sin embargo, la generación de la línea púrpura en los años 2015 en adelante impacta de forma que la mujer tuvo más canales para realizar denuncias, caso

Figura 43 . Parte 1 de carta de la Evaluación del uso del Tablero de control

contrario que sucedió en el año 2020 (año de pandemia) en el que no existían los instrumentos para realizar las denuncias respectivas.

2. Manejo de armas: El no uso de armas durante los años 2010 a 2014 nos muestra que es posible que la comunidad no tuviera tanto acceso a armas, aunque aparece un aumento este después del 2015, se debe a que también los casos aumentaron; se evidencia que el uso de armas tuvo un aumento constante desde los años 2015 en adelante, con una leve disminución en el 2020 y definitivamente con bajo reporte de casos para el 2022, posiblemente esto se deba a la migración de población vulnerable venezolana.

3. Violencia en menores de edad: Es posible analizar que el género no es diciente en las cifras de violencia contra los menores y que en su mayoría estos casos no reportan el arma utilizada o contundente, danto a entender que son menores violentados de mano directa de sus progenitores, familiares y vecinos.

Esto nos muestra que el tablero diseñado **si** puede darnos una clara ubicación de los casos específicos y de manera particular sobre cualquiera de las variables que estemos interesados a analizar o que en la exploración del mismo se puedan detectar comportamientos que no sean de resonancia y que aún estén pendientes por trabajar tanto a nivel nacional, regional o municipal.

Para nuestro gremio sería de suma importancia tener este tipo de tableros de control para realizar consultas de información real sobre la violencia doméstica (intrafamiliar) en Colombia.

Aura Nelly Rodríguez Pinilla

TRABAJADORA SOCIAL

UNIVERSIDAD COLEGIO MAYOR DE CUNDINAMARCA

Figura 44. Parte 2 de carta de la Evaluación del uso del Tablero de control