



# Universidad Europea de Valencia

**Escuela de Arquitectura y politécnica**

Grado en Ciencia de Datos

Trabajo Fin de Grado de Ciencia de Datos

**Diseño y desarrollo de cuadro de mandos para  
el seguimiento y visualización de datos de  
delincuencia y discriminación**

Autor: Elena María García López

Tutor: Alejandro Perdiguero O'Leary

Valencia, junio 2024

“Para mamá y papá, que siempre creyeron mucho más en mí de lo que yo lo hacía.

Porque es más fácil ser más valiente cuando sé que estáis a mi lado.

Espero que os sintáis tan afortunados de ser mis padres como yo me siento de ser vuestra hija.”

Espero algún día llegar a ser como vosotros.

Os quiero, con el alma.

## **Resumen**

Este proyecto tiene como objetivo desarrollar un cuadro de mando visual e interactivo, que permita analizar los datos de discriminación y delitos. Este trabajo proporcionará una herramienta integral para la prevención del delito y la mejora ciudadana, dirigido a las fuerzas policiales.

Los contenidos del TFG incluirán el diseño y la implementación del cuadro de mando, que incluirán datos relacionados con diversos tipos de delitos, casos de discriminación, ubicaciones de incidentes y tendencias temporales. Las actividades implicarán la recopilación, limpieza y análisis de datos.

Cuando se finalice el proyecto, se obtendrá una herramienta que permitirá a las autoridades identificar las áreas con mayor riesgo, patrones delictivos y los grupos vulnerables, lo que facilitará la adopción de medidas preventivas específicas.

Además, el cuadro de mandos ofrecerá la opción de interactuar con los datos, junto a otras alternativas. Permitiendo así modificar la visualización aplicando filtros, obteniendo una mayor información detallada.

Al comprender mejor los patrones delictivos y las tendencias de discriminación a través del análisis de los datos, se pueden tomar decisiones estratégicas para abordar el problema de una manera efectiva.

El análisis continuo de los datos permitiría evaluar el impacto de las políticas implementadas y ajustarlas, midiendo su eficacia a largo plazo. Promoviendo la igualdad en las ciudades y mejorando la seguridad ciudadana.

Palabras clave: Cuadro de mando, análisis de datos, seguridad ciudadana, prevención del delito.

## **Abstract**

This project aims to develop a visual and interactive dashboard that allows analysis of discrimination and crime data. This work will provide a comprehensive tool for crime prevention and citizen improvement, aimed at police forces.

The contents of the TFG will include the design and implementation of the scorecard, which will include data related to various types of crimes, discrimination cases, incident locations and temporal trends. Activities will involve data collection, cleaning and analysis.

When the project is completed, a tool will be obtained that will allow authorities to identify areas with higher risk, crime patterns and vulnerable groups, which will facilitate the adoption of specific preventive measures.

In addition, the dashboard will offer the option to interact with the data, along with other alternatives. This allowing you to modify the visualization by applying filters, obtaining more detailed information.

By better understanding crime patterns and discrimination trends through data analysis, strategic decisions can be made to address the problem effectively.

The continuous analysis of data would allow us to evaluate the impact of the implemented policies and adjust them, measuring their long-term effectiveness. Promoting equality in cities and improving citizen security.

Keywords: Dashboard, data analysis, citizen security, crime prevention.

# ÍNDICE

<b>Resumen</b>	3
<b>Abstract</b>	4
<b>Introducción</b>	7
<b>Motivación</b>	7
<b>Objetivos</b>	8
<b>Contexto</b>	9
<b>Planteamiento del Problema</b>	9
<b>1. Estado del Arte</b>	11
<b>1.1 Herramientas de Visualización</b>	11
<b>1.2 Cuadrante Mágico de Gartner</b>	12
<b>1.4 Tableau</b>	15
<b>2. Metodología</b>	16
<b>2.1 Obtención de Datos</b>	17
<b>2.3 Diseño del Panel Inicial</b>	18
<b>2.4 Diseño Bocetos</b>	19
<b>3 Proceso ETL</b>	25
<b>3.1 Paso a Paso del Procedimiento Dataset Delitos.</b>	25
<b>3.2 Paso a Paso del Procedimiento Dataset “víctimas”.</b>	27
<b>4. Proceso EDA Dataset Víctimas</b>	28
<b>Punto de Análisis</b>	35
<b>5. EDA Dataset Víctimas</b>	36
<b>6. Resultados</b>	41
<b>6.1 Creación de Visualizaciones e Informes</b>	41
<b>6.2 Requisitos Comunes</b>	41
<b>6.3 Diseño del Panel</b>	42
<b>6.4 Visualizaciones en Power BI</b>	43
<b>7. Discusiones</b>	54
<b>7.1 Implicaciones y Aplicaciones Prácticas</b>	54
<b>7.2 Discusión Sobre las Implicaciones Éticas</b>	55
<b>7.3 Contribución a los Principios, Valores Democráticos y las ODS.</b>	56

<b>8. Conclusiones</b>	<b>58</b>
<b>8.1 Cumplimiento de los Objetivos Planteados</b>	<b>58</b>
<b>8.2 Cumplimiento del Diseño Original Dashboard</b>	<b>59</b>
<b>8.3 Contribuciones del Estudio</b>	<b>60</b>
<b>8.4 Limitaciones y Futuras Líneas de Investigación</b>	<b>61</b>
<b>Referencias Bibliográficas</b>	<b>62</b>
<b>Índice de Figuras</b>	<b>63</b>
<b>Índice de Tablas</b>	<b>63</b>

## **Introducción**

En este Trabajo Final de Grado (TFG) se centra en el desarrollo de un cuadro de mando, destinado a analizar datos de delincuencia y discriminación en la región de Barcelona. Con el objetivo de prevenir incidentes y mejorar la seguridad ciudadana.

En el desarrollo de este trabajo, se destaca la importancia de la visualización de los datos para dar sentido a la información disponible. Se ha analizado cómo diversos servicios y herramientas desempeñan un papel fundamental en este proceso, resaltando Power BI como una plataforma efectiva para la visualización y análisis de datos.

## **Motivación**

La sociedad actual cada vez está más conectada, gracias al avance tecnológico y a la globalización, lo que conlleva tanto ventajas como desventajas.

Una de estas desventajas, es el aumento alarmante en el índice de delitos, especialmente en los delitos de violencia. Esto ha generado una preocupación creciente en la población y los visitantes de la ciudad. La situación requiere actuar de manera inmediata, con medidas efectivas por parte de las autoridades para garantizar el bienestar de los ciudadanos.

Conscientes de esta necesidad social, surge la motivación de este proyecto: una herramienta integral de análisis de datos que permita identificar patrones y tendencias delictivas en una línea temporal.

El análisis constante de los datos permitirá evaluar la efectividad de las medidas preventivas implementadas en un futuro y su impacto en la seguridad ciudadana. Permitiendo así garantizar una mayor seguridad.

## Objetivos

De manera esquematizada, se plantean los siguientes objetivos clave del trabajo:

- **Investigación y comprensión:** Se busca estudiar y comprender los factores que influyen en la delincuencia, así como herramientas y metodologías utilizadas en el análisis de datos, con el propósito de desarrollar una solución que puede ser implementada en diferentes regiones.
- **Desarrollo de la herramienta:** Se plantea desarrollar y diseñar un cuadro de mandos interactivo y adaptable utilizando Power BI, que permita visualizar datos de delincuencia de manera clara y efectiva.
- **Visualización intuitiva:** Diseñar una representación gráfica de los datos de delincuencia que sea fácil de entender y de utilizar para los miembros del cuerpo policial. Esto implica seleccionar gráficos que reflejen la información relevante y utilizar elementos que faciliten la interpretación de los datos.

## **Contexto**

En el desarrollo de este Trabajo de Fin de Grado (TFG), es esencial establecer las bases que permitan orientar y respaldar todo el proyecto. Esta fase inicial requiere definir el problema, e identificar el público objetivo.

Dedicar tiempo y recursos a esta fase es fundamental para garantizar que se cumplen las expectativas de los usuarios finales y garantizar que el proyecto avance de manera efectiva hacia los objetivos finales.

## **Planteamiento del Problema**

En un contexto donde la tecnología no está tan presente en las entidades policiales y la tasa de criminalidad aumenta cada año más, surge la necesidad de encontrar soluciones que permitan actuar de manera eficiente ante un tema tan sensible.

Las entidades policiales se enfrentan a grandes desafíos al tratar de entender y comprender el crecimiento continuo de la criminalidad. La falta de herramientas tecnológicas especializadas y la complejidad de los datos dificultan identificar patrones delictivos, áreas de alto riesgo y la aplicación de intervenciones eficaces.

En esta situación, la necesidad de desarrollar un cuadro de mando de análisis delictivo se vuelve crítica. Este cuadro de mando se plantea como una herramienta esencial que permitirá a las entidades policiales explorar y comprender los datos de manera más eficiente. A través de una visualización clara y detallada, proporcionará información valiosa para la toma de decisiones y planificar estrategias.

El objetivo principal es aprovechar la tecnología disponible para desarrollar un cuadro de mando que no solo visualice los datos delictivos, sino que también contenga información demográfica, para poder proporcionar un contexto completo sobre la distribución y las tendencias en la región. Permitiendo así, identificar áreas críticas, y evaluar la efectividad de las medidas que se han implementado.

En resumen, ante el aumento de la tasa de criminalidad y la falta de herramientas tecnológicas adecuadas, es fundamental desarrollar un cuadro de mando que permita a las entidades actuar de manera eficiente y efectiva.

## **Público Objetivo**

La herramienta creada en este trabajo está diseñada para ser utilizada por distintas entidades gubernamentales, policías y otras organizaciones interesadas en abordar los desafíos delictivos.

La accesibilidad y la facilidad de su uso son elementos fundamentales del diseño del sistema, está destinado a ser utilizado por una variedad de usuarios, incluidos agentes de policía y analistas de la seguridad. Por lo tanto, se centra en asegurar que sea intuitiva y fácil de manejar para todos los usuarios, sin importar su nivel de habilidad técnica.

El sistema se ha generado a partir de datos específicos de Barcelona, pero su utilidad no se limita exclusivamente a entidades gubernamentales y policiales de la región. También puede ser empleada por otras organizaciones interesadas en el análisis de datos delictivos para mejorar la toma de decisiones y la planificación estratégica en diferentes áreas geográficas.

La base de datos del sistema incluirá información detallada sobre diferentes tipos de delitos e infracciones, ubicaciones, fechas, número de víctimas y otras variables relevantes. Esto permitirá obtener información valiosa para desarrollar políticas de seguridad efectivas y crear estrategias de prevención del delito.

En resumen, el sistema proporcionará a las entidades gubernamentales, policiales y otras organizaciones una herramienta poderosa para promover la seguridad ciudadana.

# 1. Estado del Arte

El análisis de datos y las tecnologías de inteligencia artificial están siendo cada vez más utilizados en diversas áreas. Estos avances permiten el desarrollo de sistemas capaces de analizar grandes volúmenes de datos y ofrecer recomendaciones valiosas para prevenir diversos tipos de situaciones o problemas.

Otro de los ámbitos que está creciendo exponencialmente es la visualización de los datos. En el siguiente apartado se analizarán las distintas áreas que influyen en este proceso.

## 1.1 Herramientas de Visualización

Las herramientas de visualización de datos son fundamentales para comprender y comunicar información de una manera visual y efectiva. Una de sus características, es la capacidad de interactuar con los paneles, permitiendo a los usuarios explorar los datos y obtener una información más detallada sobre patrones, tendencias y relaciones.

Cada vez son más las empresas que están implementando estas tecnologías para mejorar en su toma de decisiones y obtener una mayor comprensión a partir de sus datos. La visualización facilita identificar oportunidades, problemas y permite llevar un seguimiento del rendimiento en tiempo real.

Existen numerosas aplicaciones diseñadas para la visualización de los datos, sin embargo, a la hora de elegir una herramienta para el proyecto, se ha optado por utilizar el Cuadrante Mágico de Gartner como referencia, permitiendo así tomar una decisión razonada basada en las necesidades y en los objetivos de este proyecto.

## 1.2 Cuadrante Mágico de Gartner

El Cuadrante Mágico de Gartner evalúa y clasifica a las principales empresas proveedoras de tecnología del mercado mundial mediante una representación gráfica. Clasifica a las empresas en cuatro categorías distintas:

- **Líderes:** Estas empresas destacan tanto en visión como en capacidad de ejecución, y suelen ser consideradas como referentes en el mercado.
- **Visionarios:** Tienen una visión sólida y grandes ideas, pero no demuestran la capacidad de implementarlas.
- **Desafiantes:** Proveedores con una gran capacidad de ejecución, aunque pueden carecer de una visión innovadora clara. A menudo, son competidores fuertes en el mercado, aunque pueden no ser líderes en innovación.
- **Actores de nicho:** tienen un enfoque fuerte en un área en particular en el mercado, pero pueden carecer de una visión amplia o de una ejecución firme.

El Cuadrante Mágico de Gartner se basa en un análisis detallado sustentado por una metodología bien estructurada y fundamentada en hechos.

Este gráfico se construye de dos ejes:

- **Integridad de visión:** refleja la capacidad de innovación, analizando si está siendo líder o sigue simplemente al mercado, y si la visión del desarrollo del mercado coincide con el enfoque de Gartner.
- **Capacidad de ejecución:** Considera elementos como la salud financiera del proveedor, su habilidad de responder las demandas del mercado, el progreso en el desarrollo de los productos, la eficiencia de los canales de venta y la cantidad de clientes.

Los criterios de calificación se ajustan dependiendo el mercado, evaluándolos según su importancia en una escala de “alta”, “media” o “baja”.

El Cuadrante Mágico de Gartner para plataformas Analíticas y de Business Intelligence (ABI) evalúa 20 proveedores de software distintos basándose en doce requisitos, que toda plataforma debe cumplir, estas exigencias son:

- Información automatizada.
- Preparación de los datos.
- Conectividad de fuente de los datos.
- Narración de datos.
- Visualización de datos.
- Gobernanza de los datos.
- Consulta en lenguaje natural.
- Elaboración de informes.
- Integración de Ciencia de Datos.
- Almacén de métricas.

Gartner destaca el gran alcance que ha tenido el mercado de Power BI y la posibilidad de integrarse con las otras plataformas Microsoft. Sin embargo, destaca que las herramientas Tableau y Qlik son también muy buenas opciones.

**Figura 1**



### 1.3 Power BI

En primer lugar, Power BI es una herramienta de visualización de datos muy potente, que permite crear informes interactivos y dinámicos. Esencial para el objetivo principal, ofrecer un cuadro de mando efectivo para prevenir y abordar los delitos en cada región.

Power BI tiene la capacidad de generar visualizaciones intuitivas y atractivas que permiten facilitar su comprensión y el análisis de los datos.

Entre todas las cualidades de Power BI, destacan las siguientes:

- **Innovación continua y desarrollo de la tecnología:** la empresa ha demostrado un compromiso continuo, invirtiendo en investigación y desarrollando tecnologías en diversas áreas. En particular Power BI, es un ejemplo de ese compromiso, ya que ofrece innovación en el ámbito del análisis de datos.
- **Facilidad de uso:** Microsoft ha tenido una buena adaptación en el mercado gracias a la facilidad de uso que ofrece. Power BI, en particular, como se ha comentado anteriormente, destaca por su interfaz intuitiva y su capacidad para permitir a los usuarios crear visualizaciones de datos atractivas y significativas con facilidad.
- **Conexiones de orígenes de datos:** En el caso de Power BI, esta aplicación permite conectarse desde archivos sueltos hasta plataformas online. Además, permite cargar los datos desde distintos archivos. Como txt, csv y Excel.

En resumen, Power BI se considera un líder destacado en el mercado tecnológico y de análisis de datos debido a sus servicios integrados, su compromiso con la innovación, su flexibilidad y su facilidad de uso.

## 1.4 Tableau

Dentro del marco histórico de la visualización de datos, Tableau se convierte en otra de las opciones más importantes dentro del mercado.

Es considerado una de las herramientas más influyentes en el desarrollo y la evolución del análisis de los datos. Esta herramienta permite que sea más fácil explorar y administrar los datos.

Incluso los usuarios que no tienen conocimientos técnicos pueden crear paneles y hojas de trabajo, ya que es una herramienta intuitiva y versátil.

Algunas de sus características más notables de Tableau son:

- **Facilidad de uso:** Destaca por tener una interfaz intuitiva, que permite a los usuarios explorar y visualizar los datos de una manera rápida y sencilla.
- **Conexión con múltiples fuentes de datos:** Tiene la capacidad de conectarse y combinar datos de varias fuentes, incluidas bases de datos, archivos locales, servidores en la nube. En la parte de Cloud, está preparado para su implementación en tres plataformas: Amazon Web Services, Microsoft Azure y Google Cloud Platform.
- **Potencial visual:** Ofrece una gran selección de visualizaciones de datos, desde los gráficos más simples hasta paneles interactivos, permitiendo así presentar la información de manera impactante y comprensible.
- **Seguridad:** Su potente sistema de permisos garantiza seguridad y el acceso a los datos de la empresa. Tableau dedica una gran parte de sus recursos a mejorar la experiencia de los usuarios.

## 2. Metodología

Estas herramientas son de particular importancia para las entidades gubernamentales, responsables de mantener el orden público y garantizar la seguridad ciudadana, ya que, gracias al análisis de datos y la inteligencia artificial, se pueden detectar patrones delictivos, anticipar áreas de riesgo y aplicar medidas preventivas para reducir la delincuencia y proteger la sociedad.

En este trabajo se explorarán los distintos enfoques de visualización de datos, en diferentes contextos relacionados con la criminalidad y los delitos, con el objetivo de mejorar todo su potencial para mejorar los resultados y optimizar procesos como el análisis de tendencias delictivas, identificar áreas de alto riesgo, y prevenir el crimen, utilizando Power BI como plataforma principal.

Este aspecto resalta el uso de Jupiter Notebook en los procesos de ETL y EDA, en conjunto con Power BI, para mejorar la visualización de datos y fortalecer la gestión de la seguridad pública y la criminalidad.

## 2.1 Obtención de Datos

Para crear un cuadro de mando o dashboard efectivo destinado a prevenir y abordar los delitos en cada región, es fundamental contar con los datos relevantes y recientes.

Los datos necesarios para crear este informe se han obtenido de fuentes confiables, específicamente la página oficial de los Mossos d'Esquadra, la policía autonómica de Cataluña. Esta fuente proporciona información detallada y actualizada sobre delitos, incidentes de discriminación, y otras variables relevantes.

Para la obtención de los datos, se tienen que descargar dos archivos CSV de la página web oficial de los Mossos d'Esquadra, con toda la información detallada sobre delitos, en un formato estructurado y organizado.

Una vez que los datos se han recopilado adecuadamente, se lleva a cabo un proceso de ELT y EDA. Durante este proceso, se explorarán los datos para identificar patrones y tendencias. Incluyendo la limpieza de datos para garantizar su calidad, la transformación de los datos según fuera necesario y la creación de medidas necesarias.

(Dades Obertes, s. f.)

## 2.3 Diseño del Panel Inicial

Para prevenir y abordar de manera efectiva el problema, es necesario entender los datos. En este proyecto, creamos un cuadro de mando para ayudarnos a ver y comprender los datos de forma clara y precisa.

En este proyecto, se ha empleado una metodología detallada, que combina la creatividad en el diseño y la precisión en el análisis de datos. Para poder lograr esto, se han utilizado bocetos elaborados con Excalidraw, permitiendo así una visualización previa de la estructura y su disposición. Estos bocetos, sirvieron como base para establecer las reglas del diseño, garantizando así la armonía visual.

Además, se formularon preguntas clave para cada pantalla del cuadro de mando, lo que ayudó en la toma de decisiones. Cada pregunta formulada fue seleccionada de forma cuidadosa para ofrecer una visión completa de la situación. Estas preguntas incluyeron:

- **Primera pantalla:** Primera pantalla: “¿Cuántos delitos se cometen?, ¿dónde y qué tipo son?” Esta pregunta incluye la necesidad de visualizar la cantidad total de delitos, su distribución geográfica y la clasificación por tipo de delito.
- **Segunda pantalla:** “¿Hay alguna relación entre la edad, el género, y el delito? Esta pregunta busca explorar los posibles patrones o correlaciones entre edad, género y tipo de delito cometido.
- **Tercera pantalla:** “¿Cuántos casos se han resuelto, y cuántos se han detenido? Esta pregunta se centra en evaluar el éxito de las investigaciones y el porcentaje de casos que han sido resueltos en comparación con los casos totales reportados.

Estas preguntas clave ayudaron a elegir el diseño de cada pantalla del cuadro de mando, garantizando que las visualizaciones fuesen adecuadas, informativas y prácticas para los usuarios finales.

## 2.4 Diseño Bocetos

El panel ha necesitado de tres diseños, uno para cada pregunta. En cada uno de estos diseños se ha priorizado tener un encabezado y un espacio de filtrado en la parte superior, para que el usuario pueda filtrar y explorar el cuadro de mando.

En la Figura 2 podemos observar el diseño de la primera página del cuadro de mando. Esta primera página contesta la pregunta inicial: “¿Cuántos delitos se cometen?, ¿dónde y qué tipo son?”

El diseño está formado por varios elementos clave. En la izquierda destaca un mapa interactivo permitiendo tener una visión geográfica de los incidentes y delitos.

En el centro destaca un mapa que proporciona una visión del tipo de infracción que se comete, y en la parte inferior del centro mediante un gráfico de anillos, se verá el canal mediante el cual se producen los delitos.

En el lado derecho de la página, se presentan tarjetas informativas que resumen el número total de los delitos que se han registrado. Ayudarán a tener una visión general rápida de las áreas con mayor actividad e incidencia delictiva.

Justo debajo de estas tarjetas, se encuentra el último gráfico de la página: un gráfico de nube. En esta visualización, se pretende identificar con mayor facilidad y de manera rápida los lugares donde ocurren más delitos.

Esta página del informe ofrece una visión completa y detallada de la situación delictiva, permitiendo visualizar los datos de una manera efectiva.

Figura 2

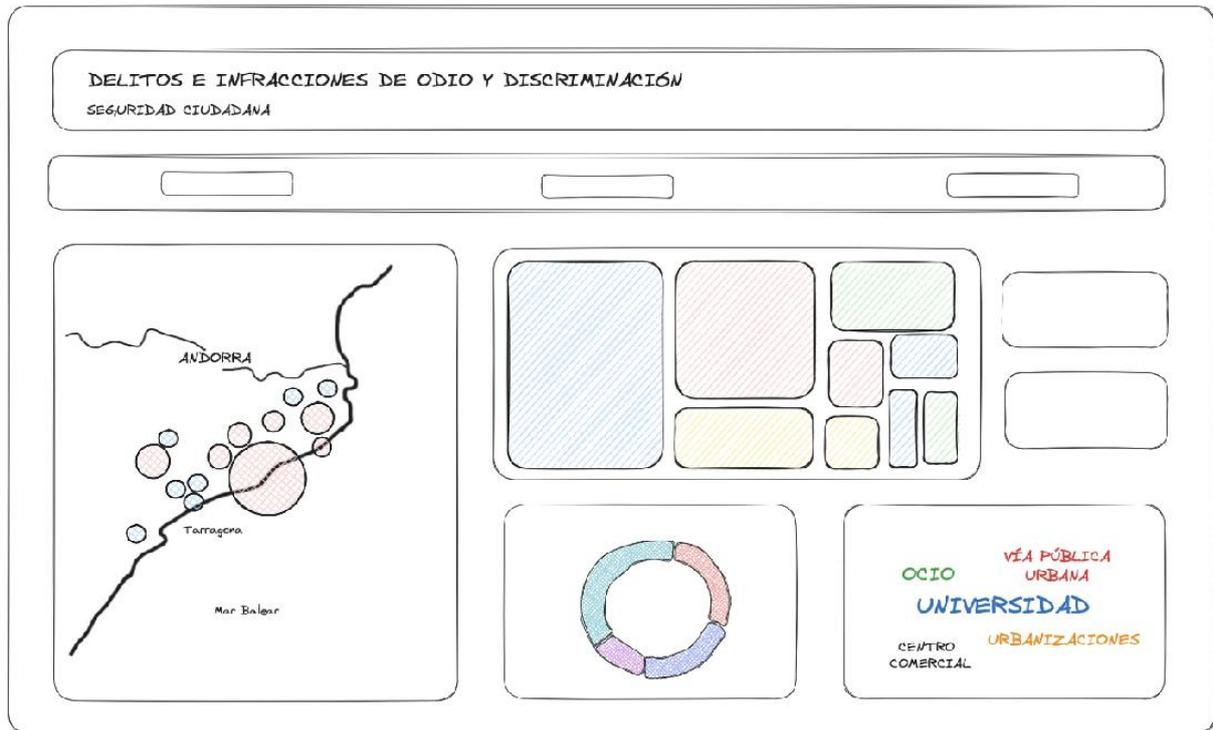


Figura 3



En la Figura 3 se observa el diseño de la segunda pantalla del cuadro de mandos. Esta pantalla contesta a la siguiente pregunta: ¿Hay alguna relación entre la edad, el género, y el delito?

En este dashboard encontramos una disposición distinta y varios elementos gráficos que permitirán responder a la pregunta.

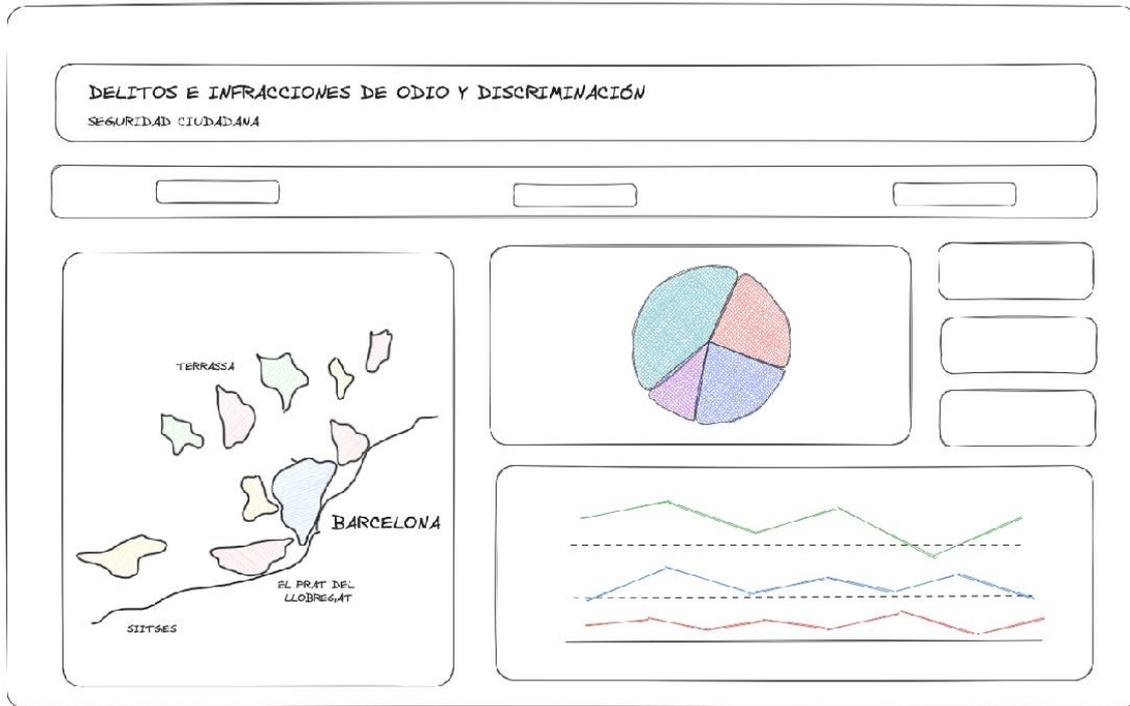
A la izquierda de la pantalla, encontramos dos gráficos de tarjeta que permiten visualizar la edad promedio tanto del autor como la de la víctima de los delitos.

Justo al lado de estos gráficos, nos encontramos con un gráfico circular que representará el porcentaje de víctimas y de autores del delito. Este gráfico permitirá visualizar y comparar visualmente la proporción.

Y en la parte derecha, podemos observar un gráfico de barras horizontales apiladas. Este gráfico ayudará a entender mejor los delitos. Además, al ser un gráfico de barras apiladas se compararán diferentes categorías, en este caso el género de las personas.

En la parte inferior izquierda de la pantalla, se encuentra un mapa que permite visualizar claramente el ámbito por el cual es producido el delito. Y en su derecha, se encuentra el último gráfico de esta pantalla. Es un gráfico de áreas que representará el total de los delitos a lo largo de los meses. Este gráfico permitirá entender ciertos patrones temporales.

**Figura 4**



Por último, en la Figura 4 representará el tercer cuadro de mando. Este responderá a la pregunta: ¿Cuántos casos se han resuelto, y cuántas personas han sido detenidas? Esta pantalla de cuadro de mando permitirá medir el rendimiento de las aplicaciones de la ley y el sistema judicial a lo que se refiere en temas de resolución de delitos.

En el boceto del cuadro de mando se ve que se incluyen elementos visuales que permiten establecer relaciones.

El foco de atención del cuadro de mando se desvía hacia la izquierda, se puede observar un mapa coroplético. En el centro del cuadro de mando se disponen dos gráficos. El gráfico situado en la parte superior permitirá comparar el porcentaje de los delitos que han sido resueltos y los delitos que ha habido detenciones.

Sin embargo, el gráfico que se encuentra en la parte inferior permitirá ver la evolución a través de los años, teniendo así una visión más amplia y poder hacer un estudio significativo de tendencias.

Por último, en la parte izquierda de la pantalla incluirá gráficos de tarjeta que permitirán hacer una comparación entre el número de casos resueltos, el número de casos que se han realizado detenciones y el número total de delitos.

## 2.5 Desglose de Datos

Los datos que se han utilizado para la creación del dashboard se representan en la siguiente tabla:

**Tabla 1**

<b>Dato</b>	<b>Significado</b>	<b>Tipo de dato</b>
Año	Año de los datos	Número entero
Num. mes	Número del mes de los datos	Nombre
Mes	Nombre del mes de los datos	Texto
Región policial	Región policial donde se han producido los delitos	Texto
Área básica policial	Nombre del área policial	Texto
Provincia	Provincia donde se han producido los delitos	Texto
Comarca	Comarca donde se producen los delitos	Texto
Municipio	Municipio donde se han producido los hechos	Texto
Lugares delitos	Lugar donde se han producido los delitos (tipo de vía, tipo de local)	Texto
Canal delitos	Canal físico o digital donde se producen los delitos	Texto
Tipo delito	Ámbito delito o infracciones administrativas	Texto
Tipo delito (Código penal) o de infracción	Leyes relacionadas con el odio y la discriminación.	Texto
Ámbito delito	Motivación que origina el delito o clasificación del delito de odio y discriminación	Texto

*(Fets Delictius I Infraccions Administratives de L'àmbit de L'odi I Discriminació | Dades Obertes de Catalunya, 2024)*

A continuación, se detallará el segundo dataset:

**Tabla 2**

<b>Dato</b>	<b>Significado</b>	<b>Tipo de dato</b>
Año	Año en el que se han producido los hechos	Nombre
Trimestre	Trimestre en el que se han producido los casos	Texto
Tipología	Tipo de violencia que sufren las víctimas	Texto
Ámbito	Grupo de personas a las cuales les afecta	Texto
Edad	Rango de edades	Texto
Víctimas atendidas	Número de mujeres que aparecen como víctimas, por violencia machista en el ámbito social o comunitario	Nombre
Agresores	N.º de agresores	Nombre

*(Víctimas Violència Masclista – Àmbit Social o Comunitari | Dades Obertes de Catalunya, 2024)*

## 3 Proceso ETL

A continuación, se presenta paso por paso el procedimiento que se ha llevado a cabo para este proyecto, utilizando el proceso ETL con Jupyter Notebook para la carga y transformación de los datasets.

El proceso ETL es una metodología que permite integrar datos de múltiples fuentes, además permite preparar los datos para el análisis y la toma de decisiones. Este proceso consta de tres fases principales: Extracción, Transformación y Carga.

### 3.1 Paso a Paso del Procedimiento Dataset Delitos.

#### **Paso 1: Extracción (Importar Librerías y Cargar Datos)**

Primero, importamos las librerías necesarias y cargamos los datasets desde archivos CSV.

#### **Descripción del Paso 1:**

Importación de Librerías y carga de Datasets: los datos son extraídos de dos archivos CSV "delitos.csv" y cargados en dataframes de Pandas.

#### **Paso 2: Transformar Datos**

En esta fase se realizan varias transformaciones en los datasets: convertir a minúsculas, eliminar acentos, reemplazar palabras con significados iguales pero escritas diferentes y eliminar duplicados.

## **Descripción del Paso 2:**

Convertir a Minúsculas: Se convierten todas las cadenas de texto a minúsculas de las columnas "Mes", "Comarca", "Tipus de lloc dels fets", "Títol del fet (Codi Penal) o Llei de referencia", "Tipus de fet".

Eliminar acentos: se utiliza la función "remove\_accents" para quitar los acentos de las cadenas de texto, de las columnas: "Comarca" y "Títol del fet (Codi Penal) o Llei de referencia".

Reemplazar palabras con significados iguales: aplicamos la función de reemplazo en las columnas "Ambit fet", "Tipus de fet", "Canal dels fets" y "Tipus de lloc dels fets", para uniformizar los términos equivalentes escritos de una manera diferente.

Eliminar Duplicados: Eliminamos las filas duplicadas para asegurar la integridad de los datos.

## **Paso 3: Cargar Datos**

Finalmente, cargamos el dataset en un nuevo archivo CSV.

Para ello se utiliza la función "to\_csv" de Pandas para guardar el dataset en un nuevo archivo CSV.

### **3.2 Paso a Paso del Procedimiento Dataset “víctimas”.**

#### **Paso 1: Extracción (Importar Librerías y Cargar Datos)**

Se muestra el dataset, la carga se ha realizado en el paso anterior.

#### **Paso 2: Transformar Datos**

En esta fase también se realizan varias transformaciones en los datasets: reemplazar palabras con significados iguales pero escritas diferentes y eliminar duplicados.

#### **Descripción del Paso 2:**

Reemplazar palabras con significados iguales y aplicamos la función de reemplazo.

Estos cambios aseguran una mejor comprensión para su posterior análisis, agrupando las edades en categorías más claras y definidas.

Verificar y eliminar duplicados: Se realiza una comprobación para identificar si existen registros duplicados en el dataset. Tras la comprobación, se confirma que no existen duplicados en los datos. Evitando así cualquier redundancia que podría causar problemas en el resultado del análisis.

#### **Paso 3: Cargar Datos**

Finalmente, cargamos el dataset en un nuevo archivo CSV.

Para ello se utiliza la función “to\_csv” de Pandas para guardar el dataset en un nuevo archivo CSV.

## 4. Proceso EDA Dataset Víctimas

El Análisis Exploratorio de Datos (EDA) es fundamental en la toma de decisiones basadas en datos. Permite explorar, entender y visualizar patrones y relaciones dentro de los conjuntos de datos.

En este proyecto, se hará un estudio distinto para cada conjunto de datos, empezando por el dataset "Víctimas". Se utilizarán herramientas como Python y bibliotecas especializadas como Pandas, Plotly.express, Seaborn. Se explorará detalladamente características como la distribución de edades, y cualquier otro aspecto relevante que permita obtener una comprensión significativa.

### **Paso 1: Visualizamos las estadísticas básicas del conjunto de datos "Víctimas".**

#### **Descripción del Paso 1:**

Columna: "Any"

Los datos abarcan desde el año 2011 hasta el año 2023. El promedio de los años se encuentra en el 2017, es decir, la mayoría de los datos se centran en los últimos años.

Esto puede tener dos significados:

- La recopilación de los datos ha mejorado.
- La incidencia de los delitos ha cambiado con el tiempo.

Columna: "Trimestre"

El tercer trimestre (3r) es el periodo más común en el que se registran los datos, con 94 incidencias.

"Esto puede significar patrones estacionales" Por tanto se tiene que estudiar si ciertos trimestres coinciden con eventos específicos o periodos de vacaciones que podrían estar influyendo.

Columna: "Tipologia"

Todos los datos se refieren a "Violencia masculista", esto nos indica que el dataset se centra específicamente en este tipo de delitos. Permitiendo así un análisis más profundo y especializado sobre la violencia de género.

Enfocarse específicamente en estos datos permitiría ayudar a diseñar estrategias de prevención y apoyo para las víctimas de violencia de género.

Columna: "Ámbit"

La edad de las víctimas varía con 8 valores únicos.

"< 20 anys" es el grupo de edad más común con 78 incidencias.

Esto indica que los menores de 20 años son los más afectados.

La distribución de los rangos de edad permite identificar los grupos de riesgo y diseñar intervenciones específicas para las distintas edades.

## **Paso 2: Ajuste de tipos de variables**

## **Paso 3: Detección y tratamiento datos ausentes**

En los datos no se observan valores nulos, por tanto, no hay que hacer ningún tratamiento de los datos.

## **Paso 4: Identificación de valores atípicos**

No se identificaron valores atípicos en la primera fase del análisis exploratorio, según lo que se ha observado en el resumen general del DataFrame obtenido mediante el método ".info()". Todos los datos parecen estar correctamente y dentro de los rangos esperados.

Es por ello, que no se ha hecho la identificación y el tratamiento de los valores atípicos.

## Paso 5: Correlación de Variables

Visualizar la distribución de las víctimas a lo largo de los años.

**Figura 5**

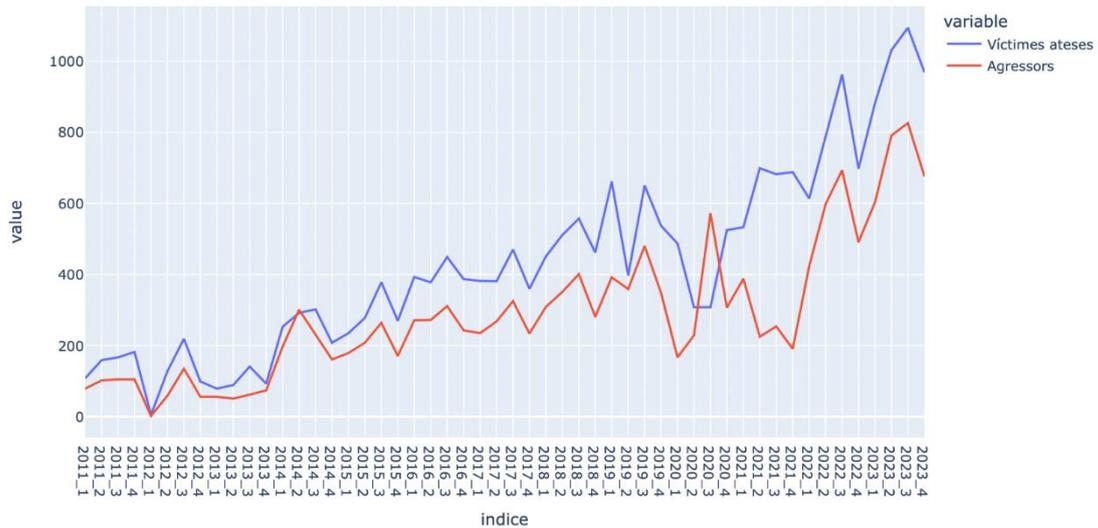


Se observa en la Figura 5 una evolución en el número de víctimas a lo largo de los años. Sin embargo, se puede observar que hay un pico hacia abajo en el 2020, posiblemente por la cuarentena.

Es por ello, que en el siguiente gráfico se estudiará los datos por año y trimestres para verificar si esta observación corresponde a esa causa.

Visualizar la distribución de las víctimas y los agresores a lo largo de los años y los cuatrimestres.

**Figura 6**



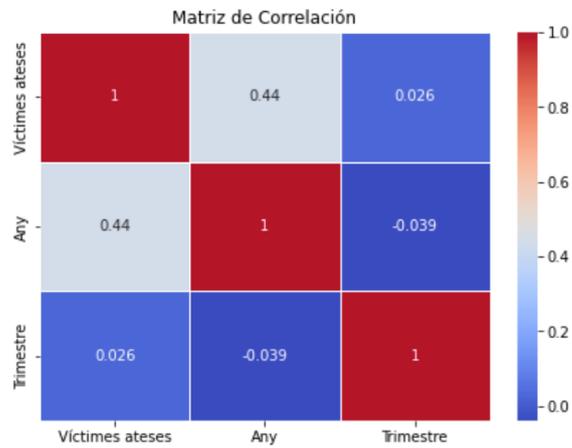
Visualizando la Figura 6 se puede afirmar la teoría. El número de víctimas y agresores disminuyó durante el periodo de cuarentena, específicamente en los dos primeros cuatrimestres del año 2020. Sin embargo, se observa un pico en el número de agresores en el tercer cuatrimestre de 2020. Esto puede ser, debido a las nuevas medidas de incorporación a la vida cotidiana, como la salida a la calle, los eventos, festivales.

Además, al observar los datos y este comportamiento, se puede concluir que la mayoría de los delitos de violencia machista ocurren en la calle.

Es por ello, que se deben implementar medidas de inmediato para prevenir estos sucesos.

## Se analiza la correlación

Figura 7



La correlación entre “víctimes ateses” y “any” es 0.44. Esto significa que hay una correlación positiva, entre el número de víctimas y el año.

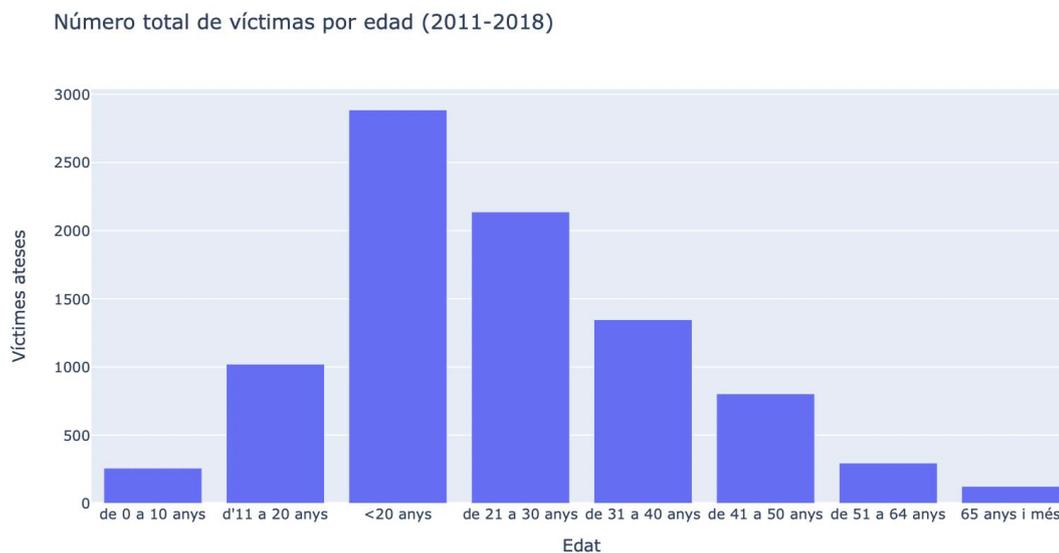
Esto podría indicar, a rasgos generales, que, a lo largo de los años, podría haber una tendencia o patrón de aumento o disminución en el número de víctimas.

Las otras correlaciones no son muy fuertes lo que significa:

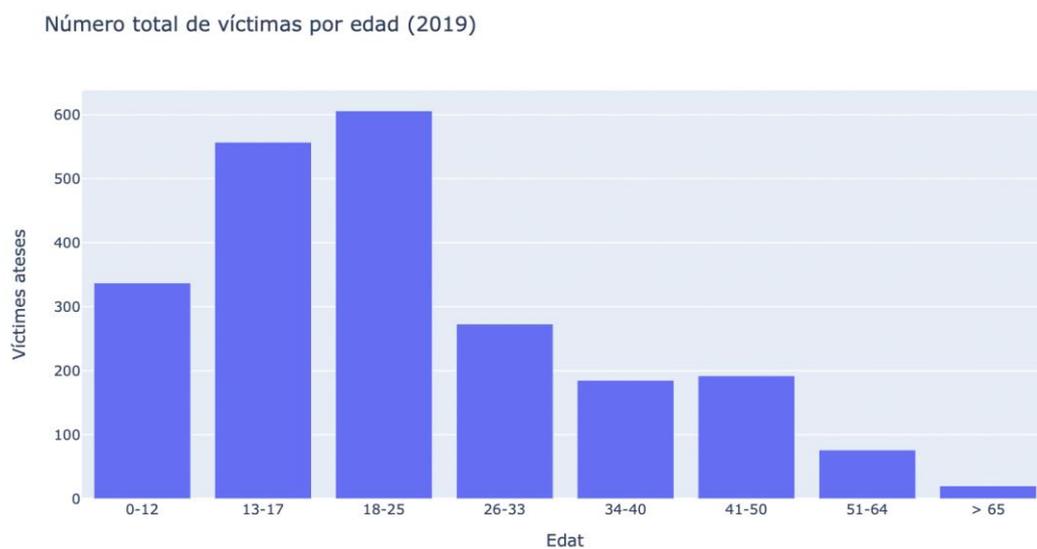
- El número de víctimas no varía según el trimestre.
- Los datos tampoco muestran relación entre el año y el trimestre en cuanto al número de víctimas.

Se visualiza el rango de edad con mayor número de víctimas

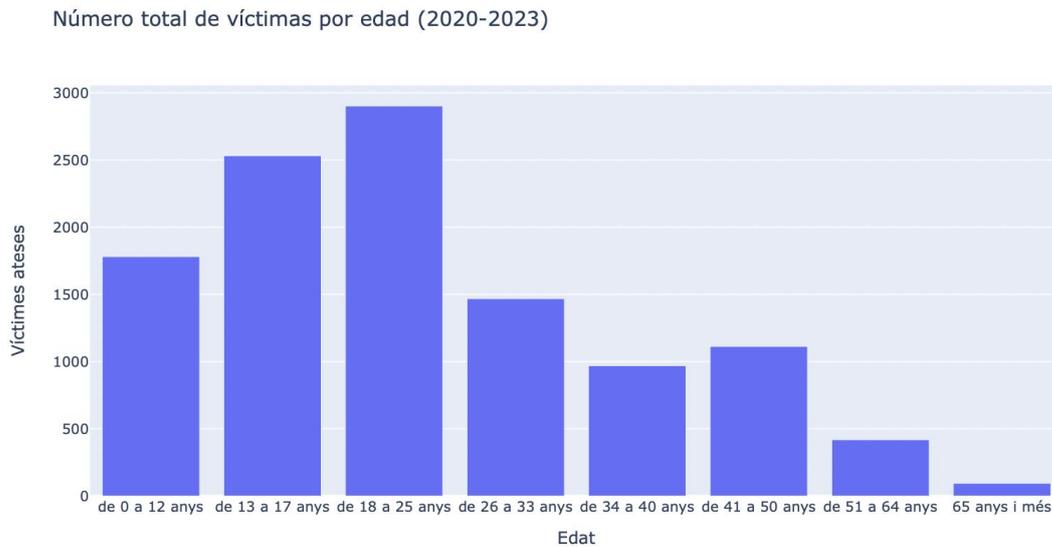
**Figura 8**



**Figura 9**



**Figura 10**



Al estudiar las Figuras 8, 9 y 10 se observa que las personas menores de 20 años representan la mayoría de las víctimas de acoso. Estos resultados subrayan la importancia de implementar estrategias de prevención y apoyo.

Identificar los factores de riesgo y la intervención adecuada puede ser un punto importante para reducir la incidencia de acoso.

Al analizar los datos de víctimas se observa una preocupante incidencia de casos de acoso en el grupo de edad de 0 a 17 años. Esta franja de edad, especialmente vulnerable, muestra cifras elevadas a lo largo de los años.

Este análisis resalta la urgencia de implementar estrategias para proteger a los niños y adolescentes de situaciones de acoso, tanto en entornos físicos, como virtuales.

Para ello se pueden implementar distintas medidas.

Participación de los padres y la comunidad: involucrar a los padres y a la comunidad de prevención del acoso, donde se debe fomentar el respeto a la comunicación abierta y la colaboración en la detección y respuesta ante el acoso.

## **Punto de Análisis**

Establecer sistemas efectivos de monitoreo y evolución para medir el impacto de las intervenciones y ajustar las estrategias.

Para ello se pueden implementar evaluaciones de impacto en programas educativos.

Paneles de control que resalten alertas sobre los aumentos repentinos en los casos de violencia machista, detectadas a través del análisis de datos.

Mediante gráficos de barras que muestren la cantidad de casos reportados en un período específico.

Sería importante valorar el desglose por tipo de acoso (verbal, físico, ciberacoso, etc).

## 5. EDA Dataset Víctimas

### Paso 1: Estudio de los Datos

Se visualizan las estadísticas básicas del conjunto de datos “Delitos” y se analizan cada una de las columnas y sus estadísticas.

#### Columna “Any”

Se observan las filas min y max para sacar la siguiente conclusión: los datos se extienden desde el año 2020 hasta el 2023.

Esto indica que los datos proporcionan una visión actualizada de los incidentes.

#### Columna “Núm. Mes”

Se observan también las filas min y max, y se saca la conclusión de que los valores van de 1 a 12, representando los meses del año.

Esto puede estar relacionado con algunos factores estacionales, como el aumento de las vacaciones escolares, las actividades al aire libre, el aumento de turismo, los festivales... teniendo así una mayor interacción social, y por ello un aumento en los delitos.

#### Columna “Mes”

Hay 12 valores únicos que representan cada mes del año. Julio es el más frecuente con 241 apariciones.

#### Columna “Región Policial (RP)”

Hay 11 regiones policiales, y RP Metropolitana Barcelona es la más común con 618 apariciones.

#### Columna “Àrea Bàsica Policial (ABP)”

En el dataset se muestran 88 Áreas policiales, siendo ABP Eixample la más común con 118 apariciones.

#### Columna "Província"

Hay 4 provincias diferentes, siendo Barcelona la más común con 1646 apariciones.

#### Columna "Comarca"

Hay 44 comarcas diferentes, siendo Barcelones la más común con 772 apariciones.

#### Columna "Tipus de lloc dels fets"

Hay 207 tipos de lugares diferentes, siendo vía pública urbana con 880 apariciones.

Esto indica la necesidad de aumentar la presencia policial en ciertas áreas públicas, para prevenir los delitos.

#### Columna "Canal dels fets"

Hay dos canales diferentes, el más común es presencial con 2056 apariciones.

#### Columna "Tipus de fet"

Hay tres tipos de hechos diferentes, siendo "delictes" el más común con 2069 apariciones.

#### Columna "Àmbit fet"

Hay 12 ámbitos diferentes, siendo LGBTI-fobia el más común con 882 apariciones.

Esto refleja una preocupación por los delitos de odio y discriminación. Esta tendencia podría indicar la necesidad de implementar políticas públicas que promuevan la igualdad y protejan los derechos fundamentales de las personas.

Además, sería interesante crear una mayor concienciación, implementando programas educativos en las escuelas. Estas charlas se podrían enfocar de distintas formas. Por ejemplo:

- Promover la inclusión
- El respeto a la diversidad
- Prevención del odio y la discriminación

Integrar estos temas ayudaría a fomentar un ambiente más seguro y comprensivo.

Columna "Nombre víctimas"

La mayoría de los registros tienen una o ninguna víctima, con un valor máximo de 7 víctimas.

## **Paso 2: Ajuste de Tipos de Variables**

Conversión columnas de tipo object a string.

## **Paso 3: Detección y Tratamiento Datos Ausentes**

Anteriormente, se ha observado que en el dataset en la columna "Nombre víctimas" había muchas columnas con nulos, es por ello, que después de hacer un estudio de los datos, se decide rellenar esos valores con la media.

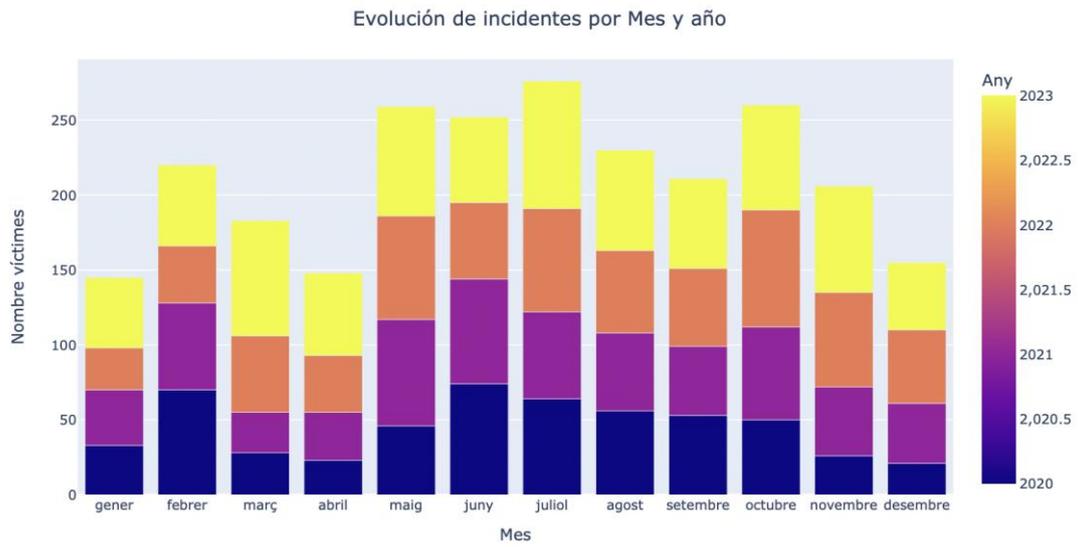
## **Paso 4: Identificación de valores atípicos**

No se identificaron valores atípicos en la primera fase del análisis exploratorio, según lo que se ha observado en el resumen general del DataFrame obtenido mediante el método ".info()". Todos los datos parecen estar correctamente y dentro de los rangos esperados.

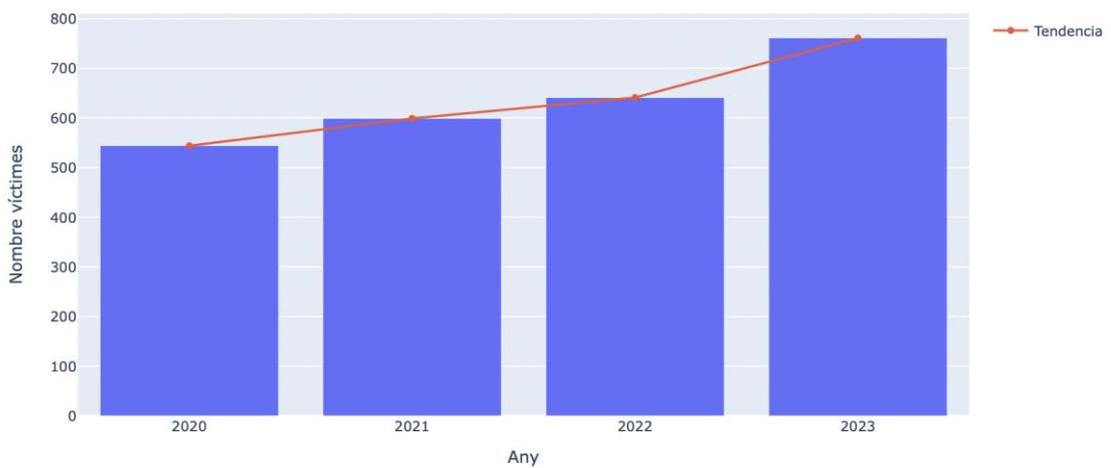
Es por ello, que no se ha hecho la identificación y el tratamiento de los valores atípicos.

## Paso 5: Correlación de Variables

### Figura 11



### Figura 12



**Figura 13**



El análisis de los datos de las víctimas muestra una tendencia de aumento en el número total de víctimas a lo largo de los años. Esta tendencia se observa en los datos recopilados y en los datos de predicciones.

Existen varios factores que pueden contribuir a disminuir el número de víctimas:

- Fomentar la denuncia a través de campañas de concienciación y proporcionando canales seguros y accesibles para que las víctimas reporten los incidentes.
- Mayor conciencia y tener un análisis continuo de los datos: esto permitirá identificar los cambios en las tendencias y evaluar la efectividad de las intervenciones implementadas.

Paso siguiente

Con estas conclusiones, se procede a realizar el estudio y la creación del dashboard.

Se realizarán más gráficos interactivos que permitan un mayor estudio de las variables y de los datos disponibles.

## **6. Resultados**

### **6.1 Creación de Visualizaciones e Informes**

Una vez se han limpiado y transformado los conjuntos de datos, el siguiente paso es visualizarlos utilizando los gráficos de Power BI. Estos informes han sido realizados de forma sencilla e intuitiva.

Los gráficos que se presentan a continuación ofrecen una visualización de los datos recopilados, identificando áreas de preocupación, tendencias e información valiosa para el desarrollo de estrategias de prevención del delito.

### **6.2 Requisitos Comunes**

Filtrado en todo el panel

Una de las características más importantes del cuadro de mando es la capacidad de filtrado dinámico. Es decir, cada vez que se selecciona un dato específico, todos los demás datos en el cuadro de mando se ajustan automáticamente.

Además del filtrado dinámico, se han implementado filtros adicionales específicos para cada pantalla. Estos filtros se han diseñado para tener una mayor precisión de visualización. Al ajustar los datos, los usuarios pueden centrarse en segmentos específicos de datos, períodos de tiempo o categorías, facilitando así identificar patrones, tendencias y relaciones clave.

Es conjunto, esta serie de características de filtrado mejoran la utilidad del informe, y permiten a los usuarios explorar y analizar los datos de manera efectiva.

### 6.3 Diseño del Panel

Una vez completados los procesos ETL y EDA, se han ajustado las preguntas clave para optimizar el estudio. Las preguntas que responderá la herramienta serán las siguientes:

- **Primera pantalla:** “Cómo se distribuyen las víctimas y agresores?”. Esta pregunta incluye la visualización del total de víctimas y agresores, así como la distribución temporal y por rangos de edad.
- **Segunda pantalla:** “¿Cuál es el perfil de la víctima y la ubicación de los delitos?”. Esta pregunta busca explorar los posibles patrones o correlaciones entre las víctimas y la ubicación del delito cometido.
- **Tercera pantalla:** “¿Cuál es el perfil demográfico?”. Aquí se centra en evaluar los delitos por Área Básica Policial y su distribución mensual.

Estas preguntas han sido perfeccionadas tras el análisis de los datos mediante ETL y EDA adaptando los gráficos y visualizaciones para proporcionar un estudio más detallado y efectivo basándonos en los datos.

Este enfoque mejorado ayudará a las entidades a tomar decisiones informadas, implementar estrategias preventivas más efectivas y mejorar la seguridad ciudadana.

## 6.4 Visualizaciones en Power BI

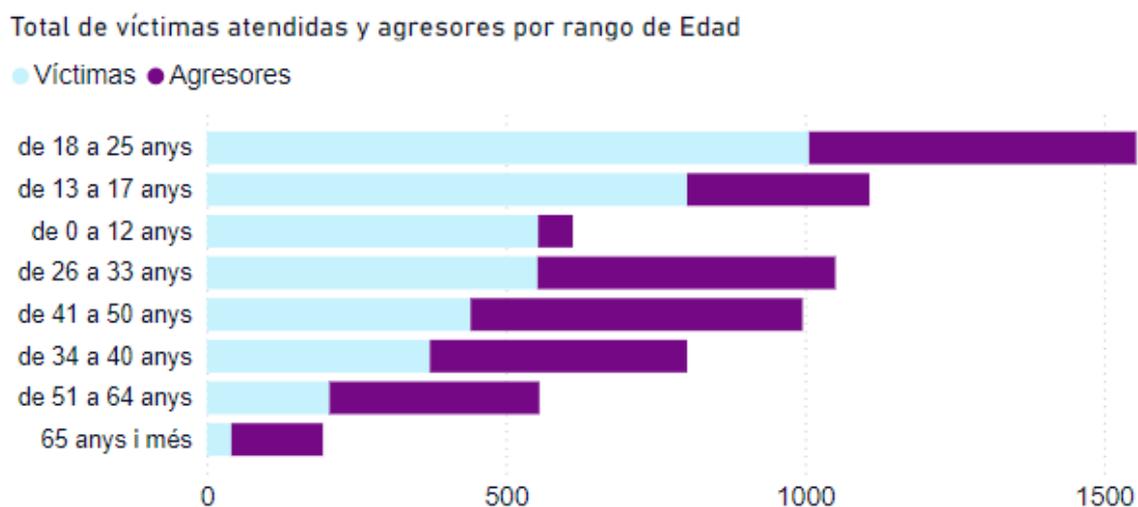
En esta sección se explorará las capacidades y aplicaciones de Power BI para la visualización de datos en el contexto del análisis de la violencia machista, los delitos y la discriminación.

Su uso en este estudio ha sido crucial para presentar la información de manera clara y accesible, facilitando la interpretación de los resultados.

El análisis de datos y su representación visual son componentes esenciales para comprender y abordar los problemas de la delincuencia y la discriminación.

En esta sección se detallará la selección de visualizaciones específicas y cómo estas herramientas han sido diseñadas para optimizar su utilidad y eficacia.

**Figura 14**



La Figura 14 muestra el total de víctimas atendidas y agresores por rango de edad en el contexto de violencia machista. Las barras horizontales representan la cantidad de víctimas (en color azul claro) y agresores (en color morado) en distintos rangos de edad.

Este gráfico permite una comparación directa entre el número de víctimas y agresores dentro de cada rango de edad, lo cual es importante para identificar las edades más afectadas y las edades de los agresores.

También ayuda a identificar los grupos de edad más vulnerables a la violencia y a la discriminación, así como los grupos de edad donde se concentra la mayor cantidad de agresores. Esta información es esencial para diseñar intervenciones específicas y dirigidas.

En la Figura 14 se muestran tendencias importantes.

Hay una alta concentración de víctimas en los rangos de edad de 18 a 25 años y 26 y 33 años, lo que indica que los jóvenes adultos son particularmente vulnerables.

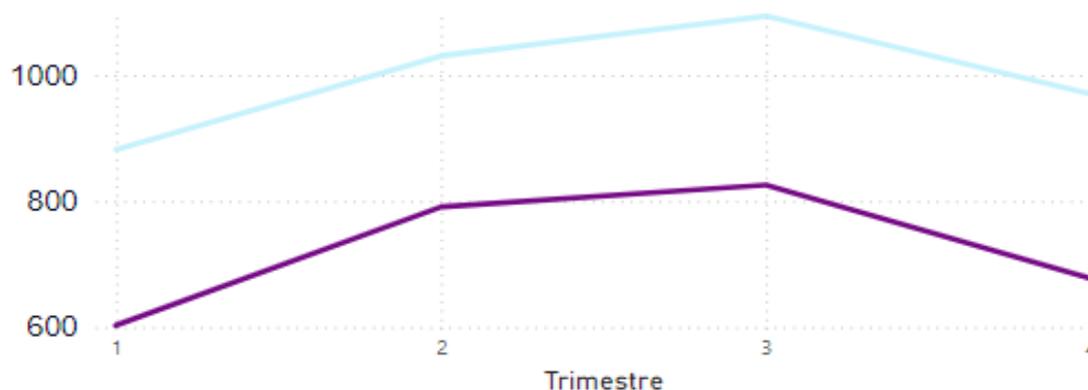
También se muestra que la mayoría de los agresores se encuentra en el mismo rango que las víctimas.

En resumen, este gráfico es una herramienta valiosa para entender la distribución de las víctimas y agresores por edad, identificar grupos de riesgo, y tomar decisiones informadas para mejorar la seguridad y reducir la violencia machista y la discriminación en la sociedad.

## Figura 15

### Víctimas y Agresores por Trimestre

● Víctimas ● Agresores



La figura 15 muestra el total de víctimas y agresores por trimestre en el contexto de la violencia machista. Las líneas representan la cantidad de víctimas y agresores (en color morado) en cada trimestre del año.

Este gráfico permite una comparación directa entre el número de víctimas y agresores a lo largo del tiempo, lo cual es importante para identificar los periodos con mayor incidencia de casos.

También ayuda a identificar tendencias y patrones temporales en la violencia machista, proporcionando información sobre como varían los casos de víctimas y agresores durante el año. Esta información es esencial para diseñar intervenciones específicas.

En la Figura 15 se muestran patrones importantes.

Se puede visualizar un aumento en la cantidad de víctimas y agresores desde el primer cuatrimestre, alcanzando su máximo en el tercer cuatrimestre. Esto significa que los casos de violencia machista aumentan a medida que avanza el año, con una concentración notable en los meses del tercer cuatrimestre.

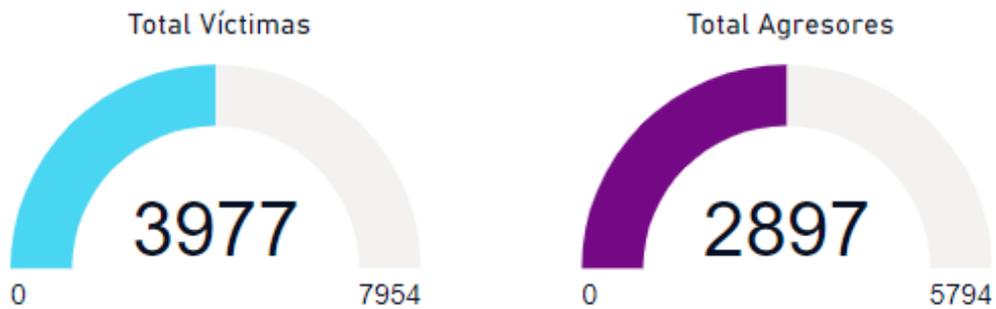
Se observa que, aunque las cifras de víctimas y agresores siguen un patrón similar, las víctimas superan el número de agresores. Esto puede reflejar casos en los que una sola persona agrede a múltiples víctimas.

Los picos en el segundo y tercer cuatrimestre indican los momentos donde se concentran más los casos de violencia machista. Esto puede asociarse a factores como las estaciones, en este caso en verano, donde las personas tienden a participar en más actividades al aire libre, eventos sociales y viajes. Esto aumenta las interacciones, lo que puede incrementar oportunidades de conflictos.

También se puede decir que las altas temperaturas y las vacaciones pueden llevar a un aumento en el consumo del alcohol, esto puede desinhibir a las personas y aumentar la probabilidad de comportamientos agresivos.

Este gráfico es una herramienta valiosa para entender la evolución temporal de las víctimas y los agresores, identificando los periodos con mayor riesgo y tomar decisiones para mejorar la seguridad y reducir la violencia en la sociedad.

**Figura 16**



La figura 16 representa dos gráficos que muestran el total de número de víctimas y agresores de violencia machista. Este tipo de gráfico permite una visualización clara del número de víctimas y agresores en un simple vistazo, lo que facilita una comprensión rápida y efectiva de los datos de estudio.

Se puede observar cifras muy altas, lo que resalta la magnitud del problema.

Este tipo de gráfico son eficientes para destacar cifras totales de una manera atractiva y fácil de entender, lo que facilita a los usuarios a identificar rápidamente la cantidad de víctimas y agresores, permitiendo una evaluación y seguimiento más eficiente de la situación.

Es importante recalcar que este tipo de gráficos es ideal para resaltar cifras clave y hacer una comparación rápida entre diferentes categorías.

**Figura 17**

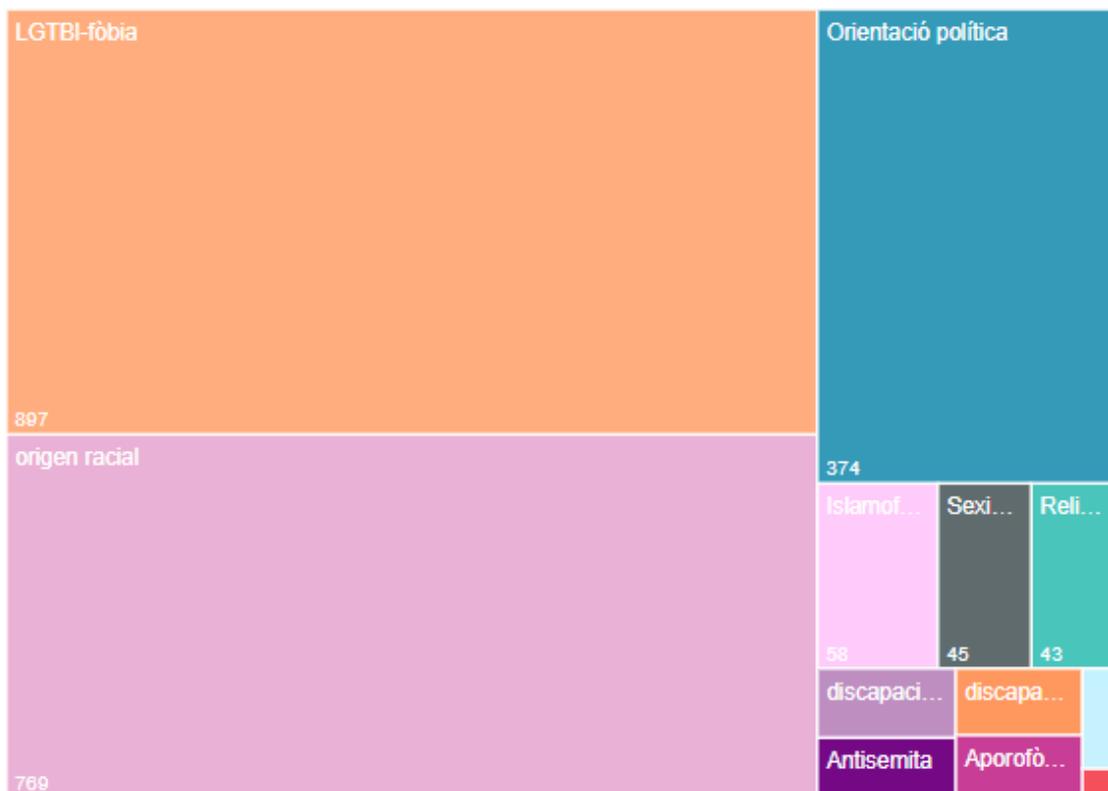
Tipo delito	Edad	Àmbito	Víctimas
Violència masclista	65 anys i més	Àmbit social/comunitari-violència sexual	41
Violència masclista	de 0 a 12 anys	Àmbit social/comunitari-violència sexual	554
Violència masclista	de 13 a 17 anys	Àmbit social/comunitari-violència sexual	803
Violència masclista	de 18 a 25 anys	Àmbit social/comunitari-violència sexual	1007
Violència masclista	de 26 a 33 anys	Àmbit social/comunitari-violència sexual	553
Violència masclista	de 34 a 40 anys	Àmbit social/comunitari-violència sexual	373
Violència masclista	de 41 a 50 anys	Àmbit social/comunitari-violència sexual	441
Violència masclista	de 51 a 64 anys	Àmbit social/comunitari-violència sexual	205
<b>Total</b>			<b>3977</b>

La figura 17 representa una tabla de datos que muestra el número de víctimas de violencia machista divididas por rango de edad y ámbito sexual. Esta tabla permite una visualización clara y organizada del número de víctimas según los distintos rangos de edad. Facilitando identificar tendencias en los datos.

Esta tabla es una herramienta eficaz para presentar y analizar los datos, ayudando así a los usuarios a identificar la distribución de las víctimas, el tipo de delito y el contexto de la violencia.

Esta tabla permitirá a los investigadores y a las autoridades comprender mejor las dinámicas y factores de riesgo asociados a cada grupo. Facilitando también el seguimiento de cambios en las tendencias, permitiendo ajustar estrategias y recursos de una manera más efectiva.

**Figura 18**



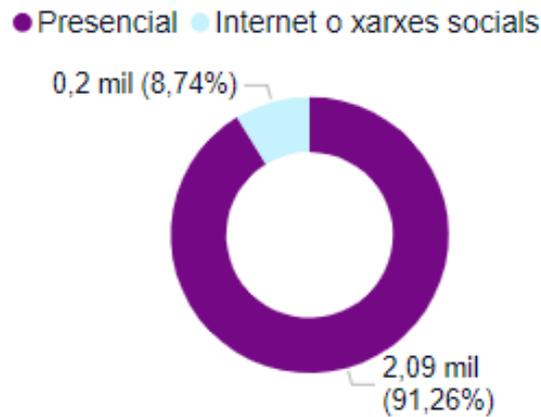
La figura 18 es un gráfico de un árbol treemap que representa datos sobre los delitos y discriminación según diversas categorías. Cada rectángulo en el gráfico corresponde a una categoría específica de discriminación, y el tamaño de cada rectángulo es proporcional al número de incidentes reportados en esa categoría.

Este gráfico es útil para visualizar rápidamente las áreas más problemáticas en términos de delitos y discriminación. Permite identificar que categorías requieren más atención y recursos para abordar los problemas de manera efectiva.

Estos resultados indican que la discriminación y los delitos de odio están distribuidos de manera desigual entre diferentes grupos, con los colectivos LGTBI, personas de diferentes orígenes raciales y aquellos con diversas orientaciones políticas siendo los más afectados.

Esto implica la necesidad de implantar políticas para proteger a estos grupos vulnerables, promover campañas de educación y concienciación para reducir los prejuicios y fomentar la inclusión.

**Figura 19**



La figura 19 representa un gráfico circular sobre la distribución de víctimas de delitos y discriminación de donde se han producido los delitos.

Este gráfico permite una visualización de la proporción entre los diferentes canales en los que se han producido los hechos. Ayuda a identificar cuál es el canal predominante donde ocurren los hechos, en este caso, el físico.

Las autoridades y organizaciones pueden usar esta información para desarrollar estrategias específicas dirigidas a prevenir y tratar los delitos y discriminación tanto en entornos físicos como digitales.

En resumen, la figura muestra que la mayoría de los hechos relacionados con delitos y discriminación ocurren en canales físicos, mientras que una menor cantidad ocurre en los canales digitales.

**Figura 20**



La figura 20 muestra una nube de palabras que representa términos relacionados con los lugares o contextos donde se han producido los hechos de delitos y discriminación.

Los términos más grandes representan las palabras más frecuentemente mencionadas en los datos, entre ellos destacan la vía pública urbana. Las palabras de un tamaño menor indican una menor frecuencia.

Esto permite identificar de manera rápida cuáles son los contextos o lugares más comunes asociados con los hechos reportados. Facilitando la asignación de recursos y la implementación de medidas preventivas en las áreas más afectadas.

Las organizaciones y las autoridades pueden usar esta información para enfocar sus esfuerzos de prevención y vigilancia en los lugares más mencionados.

En resumen, la nube de palabras indica que la mayoría de los hechos relacionados con delitos y discriminación se asocian con términos como “pública”, “vía”, y “urbana”, lo que sugiere una mayor ocurrencia en espacios públicos urbanos y transportes interurbanos.

**Figura 21**



La figura 21 representa dos indicadores numéricos (denominados “cards” en terminología de dashboards). Este tipo de gráfico es útil para mostrar cifras clave de manera directa.

Se representan dos indicadores, uno de valor fijo, que muestra un valor estático que indica el valor total de los delitos registrados, en este caso, 2289. Este indicador proporciona una referencia constante del total acumulado. Por otro lado, está el indicador dinámico, que muestra un valor que puede cambiar en respuesta a interacciones con el dashboard, como la aplicación de filtros de tiempo, tipo de delito.

Estos gráficos permiten ver rápidamente las cifras importantes sin necesidad de interpretar gráficos más complejos.

**Figura 22**

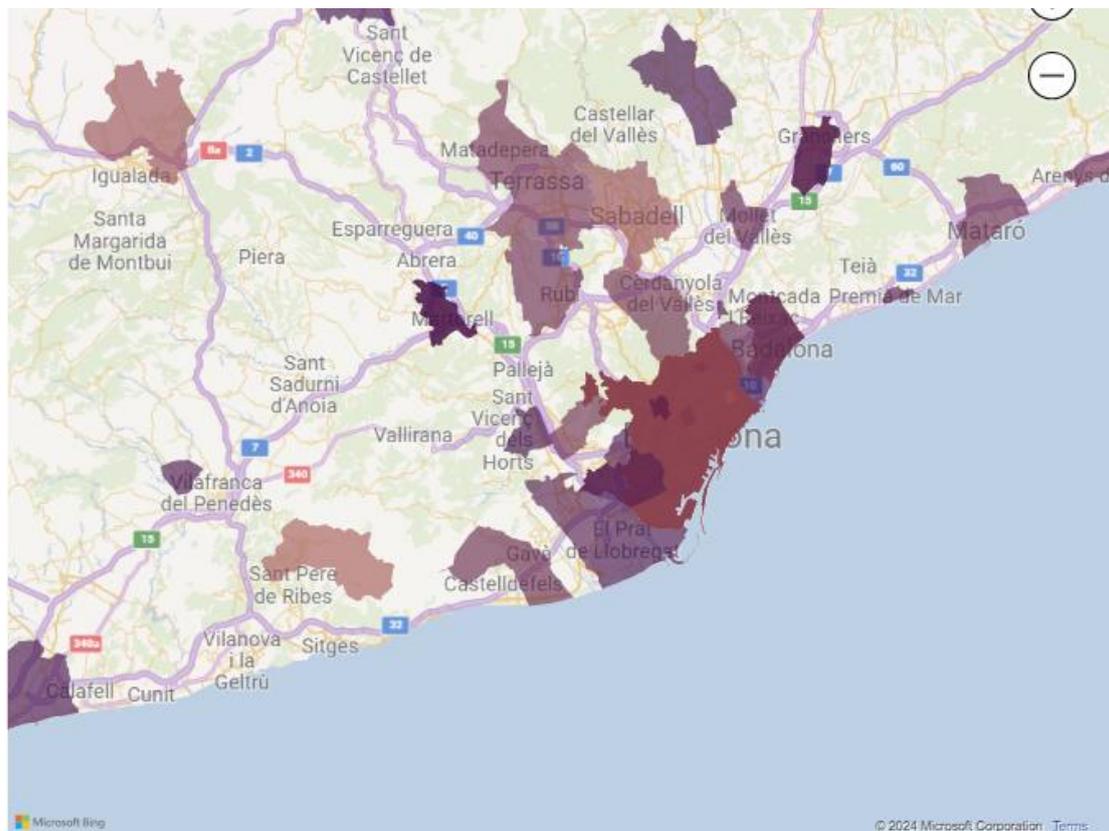


La figura 22 representa dos indicadores. En este caso la primera figura muestra el valor total de número de víctimas y la segunda el número total de incidencias. Los valores de estos gráficos se actualizan dinámicamente en función de la región seleccionada en el mapa.

Por ejemplo, si un usuario selecciona un Área Básica Policial específica, las cifras se ajustarán para mostrar el número de víctimas e incidencias en esa región.

La visualización permite comparar diferentes regiones o períodos de tiempo al observar cómo cambian las cifras cuando se interactúa con el mapa.

**Figura 23**



La figura 23 muestra un mapa de calor que representa la distribución de delitos por Áreas Básicas policiales, basado en el recuento del número de víctimas de delitos y discriminaciones.

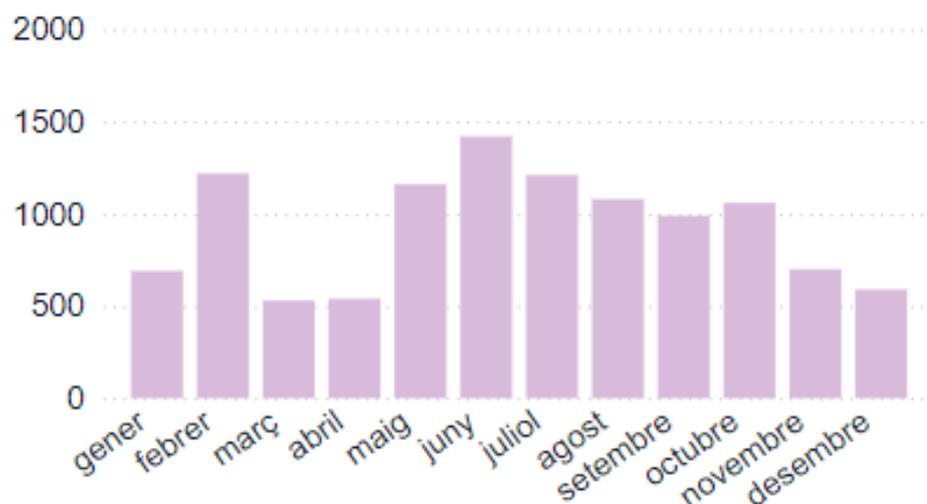
Este tipo de gráfico es útil para identificar las áreas con mayor incidencia de delitos, permitiendo aumentar la red de seguridad donde más se necesite.

Facilita la comprensión, y proporciona una forma visual e intuitiva de entender los datos, mostrando de manera clara las áreas con más víctimas.

Las zonas más conflictivas, se indican con tonos más oscuros, son aquellas con una mayor cantidad de víctimas. La zona más conflictiva es Barcelona y sus alrededores, especialmente en el centro de la ciudad. Esto se debe a varios factores.

Barcelona es un importante destino turístico, atrayendo a millones de visitantes cada año. La ciudad cuenta con numerosos eventos culturales, deportivos y sociales, que incrementan el riesgo de incidentes delictivos.

**Figura 24**



La figura 24 representa un gráfico de barras que muestra la distribución mensual del número de víctimas a lo largo de un año. Esta visualización tiene como objetivo proporcionar una visión clara de cómo varía el número de incidencias de un mes a otro.

La altura de cada barra es proporcional al número de víctimas, lo que facilita la comparación visual.

Este gráfico de barras se integra en un contexto interactivo, junto con un mapa geográfico y cifras dinámicas, lo que le permite ser aún más valioso. Ya que permite que las barras se actualicen dinámicamente según la selección de la región o filtros aplicados en la pantalla del cuadro de mando.

Al hacer clic en una barra específica, los usuarios pueden ver detalles adicionales sobre las incidencias de ese mes.

Observando el gráfico, se puede identificar meses con un mayor número de incidencias. Junio muestra el pico más alto seguido de julio y agosto.

Por tanto, comprender los patrones mensuales de incidencias puede mejorar también a las estrategias de prevención y preparación ante futuros eventos.

## 7. Discusiones

### 7.1 Implicaciones y Aplicaciones Prácticas

Las implicaciones y aplicaciones de este estudio buscan mejorar las estrategias de compromiso y concienciación sobre la violencia machista, la discriminación y los delitos. Se ha descubierto una oportunidad para que los profesionales ajusten sus enfoques y tácticas después de analizar los datos detalladamente, con el fin de comprometerse más efectivamente con la seguridad ciudadana. Entender el análisis y el estudio de los datos, junto con la visualización mediante gráficos, ayuda a crear mejoras significativas en las estrategias de prevención y respuesta.

El estudio de los patrones y tendencias de la violencia machista y la discriminación se traduce directamente en prácticas efectivas. Las organizaciones pueden colaborar directamente en la aplicación de nuevas medidas y en el uso de la herramienta dashboard para monitorear, evaluar y ajustar las intervenciones en tiempo real, formando un mayor impacto y cambio positivo en la sociedad.

Es crucial porque permite a las autoridades y a las organizaciones de la sociedad civil identificar áreas críticas que requieren atención urgente, permitiendo una asignación más eficiente de recursos y una respuesta más rápida y precisa. Además, la visualización clara de los datos mejora la comunicación y la comprensión.

En conclusión, los hallazgos de esta investigación establecen un fundamento firme para mejorar las estrategias en relación con la violencia machista y la discriminación. Comprender las percepciones y la realidad de estos problemas, permite diseñar intervenciones más efectivas y adecuadas. Es por ello, aplicar tácticas basadas en datos y adaptadas a las características específicas de cada contexto es esencial para promover un ambiente seguro y equitativo para todos los individuos.

Implementar estas estrategias no solo mejora la eficacia de las políticas públicas y las intervenciones sociales, sino que también fortalece la confianza de la comunidad en las instituciones encargadas de su protección y bienestar. Facilitando así la creación de una sociedad más justa y resiliente.

## **7.2 Discusión Sobre las Implicaciones Éticas**

La recopilación y tratamiento de datos sensibles, que abarca este trabajo sobre la violencia machista, los delitos y la discriminación, implica desafíos éticos significativos en términos de privacidad, confidencialidad y manejo responsable de la información.

La violencia machista y los delitos relacionados con la discriminación son temas delicados que involucran la vida privada y la seguridad personal de las víctimas. Es por ello, que en este trabajo se trata con extremo cuidado la recopilación y el tratamiento de los datos. Se implementan técnicas avanzadas para mejorar la visualización de los datos, asegurando que la visualización sea clara, precisa y comprensible.

Esto no solo permite evitar los errores en la interpretación de los resultados, sino que también promueve una representación justa de los datos. Ya que el manejo inadecuado de esta información podría generar malentendidos o interpretaciones erróneas sobre la magnitud de estos problemas sociales.

Los resultados obtenidos de este análisis deben ser utilizados no solo para comprender mejor los desafíos existentes, sino también para informar políticas públicas y acciones que promuevan el bienestar y la justicia social.

### **7.3 Contribución a los Principios, Valores Democráticos y las ODS.**

El diseño y desarrollo este proyecto basado en el análisis, seguimiento y visualización de los datos sobre la violencia machista, discriminación e infracciones, coincide con los principios y valores democráticos, además de contribuir a los Objetivos de Desarrollo Sostenible (ODS).

Este trabajo cumple con los principios de accesibilidad universal y diseño. Al utilizar herramientas intuitivas para mostrar los gráficos, se facilita su interpretación y comprensión. Esto permite ampliar su público, ya que se asegura que la información que se proporciona sea accesible y comprensible para una gran cantidad de personas, independientemente de sus capacidades.

Además, este estudio coincide con los Objetivos de Desarrollo Sostenible (ODS), específicamente con el ODS 5 de igualdad de Género. Al analizar y visualizar datos sobre la violencia machista, se está fomentando la igualdad de género.

El estudio de los datos y la visualización clara y detallada permite identificar áreas críticas, además de concienciar sobre la gravedad del problema y promover un cambio en la sociedad frente a la violencia de género.

Asimismo, este trabajo cumple con el ODS 10, basado en el principio de reducción de las desigualdades, tratando las diferencias sociales y políticas. El análisis de datos sobre la discriminación e infracciones administrativas en el ámbito del odio y la discriminación permite identificar los problemas que está teniendo la sociedad, e identificar las áreas donde las políticas actuales están fallando en su objetivo de promover la igualdad y justicia social.

Por ende, este análisis contribuye a generar políticas más inclusivas y efectivas que aseguren la protección de los derechos de todas las personas, independientemente su origen étnico, orientación sexual o cualquier otra característica protegida por la ley

También añadir, el cumplimiento del ODS 11 de Ciudades y Comunidades Sostenibles, Este trabajo incluye la creación de un mapa que permite identificar patrones geográficos y la evaluación de la distribución del número de víctimas. Permitiendo así el monitoreo, y diseñar intervenciones más precisas y efectivas para mejorar la seguridad y el bienestar en la comunidad, promoviendo así entornos urbanos más seguros.

En resumen, este estudio sobre el análisis y visualización de un cuadro de mandos para el seguimiento sobre datos de delincuencia y discriminación cumple con los principios éticos y responsables, además de contribuir en los Objetivos de Desarrollo Sostenible (ODS).

## **8. Conclusiones**

### **8.1 Cumplimiento de los Objetivos Planteados**

Se puede afirmar que todos los objetivos planteados se han logrado de manera satisfactoria.

Desde el principio, el estudio se propuso analizar y visualizar datos relacionados con la violencia machista, los delitos y la discriminación, con el fin de proporcionar una base sólida para la implementación de políticas públicas y estrategias de intervención.

A través de un enfoque metódico, se ha conseguido reunir y procesar una cantidad significativa de datos, los cuales han sido presentados de manera clara y accesible utilizando herramientas avanzadas de visualización como Power BI.

Los resultados obtenidos no solo han cumplido con los objetivos planteados, sino que también han proporcionado información valiosa para formular nuevas políticas y la toma de decisiones. Se han identificado áreas críticas que requieren atención urgente, y se han ofrecido recomendaciones prácticas basadas en datos detallados.

Además, los hallazgos contribuyen a una comprensión actualizada sobre la situación de dichos temas, lo cual es crucial para abordar el problema con diferentes enfoques y perspectivas.

El estudio no solo establece la base para desarrollar estrategias efectivas, sino que también introduce la herramienta del dashboard, que permite una monitorización continua y en tiempo real de las intervenciones y políticas implementadas. Este sistema de seguimiento facilita la evaluación y el ajuste de las estrategias según sea necesario, asegurando una respuesta ágil y adaptativa a los cambios y necesidades emergentes.

## **8.2 Cumplimiento del Diseño Original Dashboard**

Durante la fase de desarrollo del dashboard, se llevó a cabo una revisión de los bocetos iniciales de las pantallas. Esta revisión no solo se centró en cumplir con los requisitos iniciales de diseño, sino que también en explorar nuevas posibilidades para maximizar la usabilidad y la eficacia.

La evaluación detallada reveló que ciertos gráficos requerían de ciertos ajustes para mejorar la interpretación de los datos por parte del usuario. Como resultado se implementaron nuevas representaciones visuales que facilitaron una comprensión más clara y profunda de la información presentada.

Los cambios no solo se limitaron en aspectos estéticos, sino que también se enfocaron en optimizar la función general del dashboard. Además, estas mejoras contribuyeron en la capacidad del dashboard para apoyar procesos analíticos y facilitar la toma de decisiones.

En resumen, la evolución del diseño no solo cumplió con las expectativas iniciales, sino que también se mejoró conforme avanzaba el proceso de desarrollo. Este enfoque mejoró la efectividad de la herramienta en función de las necesidades.

### **8.3 Contribuciones del Estudio**

El análisis de los datos ha proporcionado una visión integral del análisis de incidencias y víctimas mediante un enfoque combinado ETL, EDA y visualización con Power BI. Este estudio ha demostrado múltiples beneficios y ha contribuido significativamente a la comprensión y gestión de los datos.

En primer lugar, la visualización de datos mediante gráficos de barras, nube de palabras, treemaps e indicadores numéricos en Power BI facilitó una comprensión más profunda y una mejor toma de decisiones. Estas visualizaciones no solo ofrecen una representación clara y concisa de los datos, sino que también permiten identificar rápidamente áreas de preocupación y tomar medidas correctivas adecuadas.

El proceso EDA, se incluyó un gráfico de predicciones, que destacó la importancia de tomar conciencia sobre el caso, subrayando tendencias y futuros escenarios. Este gráfico es crucial para la planificación y preparación, ya que proporciona una visión anticipada de los siguientes años.

Además, el uso de Jupyter Notebook para el análisis EDA y la limpieza de datos en el proceso ETI resultó efectivo para preparar datos precisos para la visualización. La combinación de estas metodologías ha asegurado la calidad y exactitud de los datos, lo que es esencial para el tratamiento de los datos con carácter sensible.

El estudio también ha identificado patrones claros en la distribución temporal y geográfica de las incidencias, destacando junio, julio y agosto como meses de mayor número de víctimas. Esta información es de especial ayuda para las autoridades y organizaciones, ya que permite concentrar sus esfuerzos y recursos durante los períodos críticos.

En resumen, este estudio asegura una comprensión profunda y una toma de decisiones informada. Esto es fundamental para mejorar la gestión y prevención de incidencias en cualquier ámbito.

## 8.4 Limitaciones y Futuras Líneas de Investigación

Actualmente, los datos del dashboard del mapa no se actualizan en tiempo real. La implementación de técnicas de web scraping permitiría obtener datos más actualizados y mejorar el monitoreo continuo.

Además, el dataset de víctimas carece de ciertas columnas que podrían enriquecer el análisis. La inclusión de más detalles, como el tipo de incidente, la ubicación específica, la hora del día, el tipo de agresión, podrían proporcionar una visión más completa y detallada.

Es por ello que desarrollar técnicas automatizadas de web scraping para recoger datos en tiempo real desde fuentes confiables ayudaría a mantener el dashboard actualizado y proporcionar información más precisa para el monitoreo y la toma de decisiones.

Asimismo, ampliar el dataset de víctimas con nuevas columnas que incluyan información como la clasificación detallada de tipos de incidentes, coordenadas geográficas exactas, registro de la hora específica del incidente, serían puntos clave a la hora de implantar medidas.

Estas líneas de investigación no solo mejorarían la precisión del análisis, sino que proporcionarían una herramienta más útil para la toma de decisiones en tiempo real.

## Referencias Bibliogrficas

Dades obertes. (s. f.). Mossos D'Esquadra.

[https://mossos.gencat.cat/ca/els\\_mossos\\_desquadra/indicadors\\_i\\_qualitat/dades\\_obertes/](https://mossos.gencat.cat/ca/els_mossos_desquadra/indicadors_i_qualitat/dades_obertes/)

Fets delictius i infraccions administratives de l'mbit de l'odi i discriminaci |

Dades obertes de Catalunya. (2024, 14 marzo).

[https://analisi.transparenciacatalunya.cat/Seguretat/Fets-delictius-i-infraccions-administratives-de-l-/y48r-ae59/about\\_data](https://analisi.transparenciacatalunya.cat/Seguretat/Fets-delictius-i-infraccions-administratives-de-l-/y48r-ae59/about_data)

Vctimes violncia masclista – mbit social o comunitari | Dades obertes de

Catalunya. (2024, 9 febrero).

[https://analisi.transparenciacatalunya.cat/Seguretat/V-ctimes-viol-ncia-masclista-mbit-social-o-comunit/6rcq-y46b/about\\_data](https://analisi.transparenciacatalunya.cat/Seguretat/V-ctimes-viol-ncia-masclista-mbit-social-o-comunit/6rcq-y46b/about_data)

## Índice de Figuras

Figura 1	13
Figura 2	20
Figura 3	20
Figura 4	22
Figura 5	30
Figura 6	31
Figura 7	32
Figura 8	33
Figura 9	33
Figura 10	34
Figura 11	39
Figura 12	39
Figura 13	40
Figura 14	43
Figura 15	44
Figura 16	46
Figura 17	47
Figura 18	48
Figura 19	49
Figura 20	50
Figura 21	51
Figura 22	51
Figura 23	52
Figura 24	53

## Índice de Tablas

Tabla 1	23
Tabla 2	24