



Universidad Internacional de La Rioja  
Escuela Superior de Ingeniería y Tecnología

Máster Universitario en Análisis y Visualización de Datos  
Masivos/ Visual Analytics and Big Data

## Comparativa de herramientas de visualización para el análisis de datos criminalísticos

Trabajo fin de estudio presentado por:	Susana Barrera García
Tipo de trabajo:	Comparativa de soluciones.
Director/a:	Diego Fernández García
Fecha:	Julio 2025

## Resumen

Este Trabajo Fin de Máster compara tres herramientas de visualización de datos —Tableau, Power BI y Python (Plotly/Matplotlib)— aplicadas al análisis de tendencias criminales en Europa, empleando el modelo CRISP-DM y un sistema de evaluación multicriterio propio. Se inició con la recopilación, limpieza y transformación de los datos de Eurostat para garantizar su calidad. A continuación, se diseñaron y desarrollaron visualizaciones representativas en cada plataforma, atendiendo a recomendaciones de diseño para optimizar claridad e interpretación. Finalmente, se evaluaron cinco criterios clave: facilidad de uso, calidad visual, rendimiento, flexibilidad de análisis y curva de aprendizaje.

Los resultados revelan que Power BI ofrece una implementación rápida y una interfaz intuitiva, idónea para analistas sin conocimientos de programación y entornos empresariales que requieren informes automatizables. Python destaca por su flexibilidad y capacidad de automatización, siendo la opción preferente para perfiles técnicos y académicos que necesitan personalización avanzada. Tableau se sitúa en un punto intermedio, equilibrando estética y funcionalidad para la presentación de resultados a audiencias no especializadas. En conclusión, no existe una herramienta universalmente “mejor”; la selección debe basarse en el contexto y las necesidades específicas del usuario, proporcionando esta guía una base práctica para la toma de decisiones en proyectos de visualización de datos.

**Palabras clave:** visualización de datos, análisis criminalístico, Tableau, Power BI, Python.

## Abstract

This project compares three data visualization tools—Tableau, Power BI, and Python (Plotly/Matplotlib)—applied to the analysis of criminal trends in Europe, using the CRISP-DM model and a proprietary multicriteria evaluation system. It began with the collection, cleaning, and transformation of Eurostat data to ensure its quality. Next, representative visualizations were designed and developed on each platform, following best-practice design guidelines to optimize clarity and interpretation. Finally, five key criteria were assessed: ease of use, visual quality, performance, analytical flexibility, and learning curve.

The results reveal that Power BI offers rapid deployment and an intuitive interface, making it ideal for analysts without programming knowledge and for business environments requiring automatable reports. Python stands out for its flexibility and automation capabilities, being the preferred option for technical and academic profiles needing advanced customization. Tableau occupies a middle ground, balancing aesthetics and functionality for presenting results to non-specialist audiences. In conclusion, there is no single universally “best” tool; selection should be based on the context and the user’s specific needs, and this guide provides a practical foundation for decision-making in data visualization projects.

**Keywords:** data visualization, criminal analysis, Tableau, Python, Power BI.

## Índice de contenidos

1.	Introducción .....	1
1.1.	Motivación .....	1
1.2.	Planteamiento del trabajo .....	1
1.3.	Estructura del trabajo .....	2
2.	Contexto y estado del arte .....	3
2.1.	Contexto del problema .....	3
2.2.	Estado del arte .....	4
2.2.1.	Fundamentos de la visualización.....	4
2.2.2.	Visualización en el análisis criminalístico.....	9
2.2.3.	Herramientas de visualización: antecedentes y clasificación. ....	11
2.2.4.	Comparativas previas.....	16
2.3.	Conclusiones .....	17
3.	Objetivos concretos y metodología de trabajo.....	18
3.1.	Objetivo general.....	18
3.2.	Objetivos específicos.....	18
3.3.	Metodología del trabajo .....	19
4.	Marco normativo.....	20
5.	Desarrollo específico de la contribución.....	21
5.1.	Planteamiento de la comparativa. ....	21
5.1.1.	Identificación del problema y motivación .....	21
5.1.2.	Soluciones seleccionadas y criterios de evaluación .....	21
5.2.	Desarrollo de la comparativa .....	22
5.2.1.	Preparación del entorno y carga de datos.....	22
5.2.2.	Implementación de visualizaciones comunes. ....	33

5.2.3.	Evaluación técnica de cada herramienta.....	47
5.3.	Discusión de análisis y resultados. ....	48
5.3.1.	Ventajas y desventajas observadas .....	48
5.3.2.	Casos de uso más adecuados para cada solución.....	49
5.3.3.	Consideraciones finales sobre la aplicabilidad. ....	50
6.	Código fuente y datos analizados .....	50
6.1.	Código fuente.....	50
6.2.	Datos Analizados.....	51
7.	Conclusiones.....	51
8.	Limitaciones y prospectiva .....	53
8.1.	Limitaciones .....	53
8.2.	Trabajo futuro .....	54
	Referencias bibliográficas .....	56

## Índice de figuras

Figura 1. Visualización de datos de delincuencia con ArcGIS. ....	3
Figura 2. Ejemplo de gráfico con exceso de elementos decorativos. ....	5
Figura 3. Ejemplo de gráfico con la proporción de tinta ideal. ....	6
Figura 4. Rediseño de un gráfico original del Departamento de Salud de Georgia sobre casos de COVID-19. ....	7
Figura 5. Interfaz de CompStat 2.0 del Departamento de Policía de Nueva York, que permite a los ciudadanos acceder a estadísticas delictivas actualizadas y mapas interactivos. ....	10
Figura 6. Visualización del riesgo de criminalidad en tiempo real en un barrio de Río de Janeiro mediante la plataforma CrimeRadar. ....	11
Figura 7. Ejemplo de dashboard interactivo creado con Apache Superset ....	12
Figura 8. Visión general del dataset de violencia sexual. ....	22
Figura 9. Visión general del dataset de homicidios. ....	23
Figura 10. Visión general del dataset de prisioneros. ....	24
Figura 11. Interfaz Power Query ....	24
Figura 12. Tabla dimension localización. ....	26
Figura 13. Tabla dimensión año. ....	26
Figura 14. Tabla hechos asaltos sexuales. ....	27
Figura 15. Tabla hechos prisioneros. ....	27
Figura 16. Tabla hechos homicidios. ....	27
Figura 17. Vista del modelo estrella. ....	28
Figura 18. Carga y limpieza con Python. ....	29
Figura 19. Dataset final en Python ....	30
Figura 20. Paso de limpieza en Tableau. ....	31
Figura 21. Paso de salida en Tableau ....	31

Figura 22.Tabla de prisioneros limpia en Tableau.....	32
Figura 23.Gráfica 1 de barras agrupadas en PowerBI .....	34
Figura 24.Gráfica 2 de barras agrupadas en PowerBI .....	34
Figura 25.Filtro de año en el dashboard de barras agrupadas. ....	36
Figura 26.Highlight por defecto de PowerBI. ....	35
Figura 27. Gráfico de líneas en PowerBI. ....	36
Figura 28.Mapa de burbujas en PowerBI. ....	37
Figura 29.Grafico de columnas agrupadas con Python. ....	38
Figura 30. Grafico de columnas 2 en Python.....	39
Figura 31.Gráfica 1 Python con Plotly. ....	40
Figura 32.Gráfica 2 Python con Plotly. ....	40
Figura 33.Gráfica evolución con Python. ....	41
Figura 34.Mapa de burbujas con Python. ....	41
Figura 35.Dataset con coordenadas añadidas. ....	42
Figura 36.Barras agrupadas en Tableau. ....	43
Figura 37.Barras agrupadas 2 en Tableau .....	44
Figura 38.Gráfica evolución Tableau. ....	45
Figura 39.Mapa de burbujas en Tableau.....	46
Figura 40.Variables creadas automáticamente en Tableau. ....	46

## Índice de tablas

Tabla 1.Comparación de criterios entre herramientas. ....	48
Tabla 2. Comparación de objetivos planteados con resultados obtenidos. ....	52



# 1. Introducción

## 1.1.Motivación

Se observa la necesidad de estudiar la criminalidad y la violencia en ciertas zonas por su repercusión en la calidad de vida y en la percepción de seguridad. A menudo, las causas se relacionan con factores socioeconómicos o dificultades de acceso a oportunidades, por lo que un examen profundo de estos elementos ofrece una comprensión más precisa del incremento o disminución de los índices delictivos. Asimismo, resulta pertinente comparar los hallazgos con otras regiones para identificar patrones comunes y plantear soluciones efectivas.

Por otra parte, se considera esencial abordar este fenómeno con un enfoque interdisciplinario que integre técnicas de Big Data y métodos de visualización de grandes volúmenes de información. Dicho enfoque facilita el análisis estadístico exhaustivo y la detección de tendencias que permiten diseñar medidas de intervención más ajustadas a la realidad. Estas aproximaciones, desarrolladas en el marco del máster, aportan herramientas innovadoras para profundizar en el estudio científico de la criminalidad y la violencia, promoviendo entornos más seguros y mejor informados.

## 1.2.Planteamiento del trabajo

En el ámbito de la ingeniería de software y el desarrollo web, las instituciones públicas y privadas disponen de grandes volúmenes de datos sobre criminalidad que, por su complejidad y heterogeneidad, resultan difíciles de procesar y presentar de forma accesible. Además, la necesidad de garantizar la confidencialidad y la integridad de la información impone requisitos de ciberseguridad en todo el ciclo de vida del dato. Frente a esta situación, se detecta la carencia de una solución integrada que permita extraer, limpiar y visualizar datos de criminalidad de manera ágil, reproducible y segura, optimizando la toma de decisiones basada en evidencia.

Para dar respuesta a esta necesidad, se propone el desarrollo de un estudio comparativo de tres herramientas de visualización (Power BI, Tableau y Python con Plotly/Matplotlib) aplicado a datos de Eurostat, así como el diseño de una arquitectura de prototipo de aplicación web que incorpore estas visualizaciones con controles de acceso y buenas prácticas de seguridad. Con ello, se pretende demostrar cómo puede implementarse un pipeline completo de

ingeniería de software, desde la obtención y normalización de datos hasta la publicación de dashboards interactivos que cumpla con los estándares de desarrollo web y los requisitos del GDPR.

### 1.3. Estructura del trabajo

**Capítulo 1. Introducción.** Presenta el contexto general del estudio, explica la motivación detrás de la comparación de herramientas de visualización de datos en el ámbito de la criminalidad europea y plantea de forma breve el problema detectado.

**Capítulo 2. Contexto y estado del arte.** Ofrece una visión detallada del dominio de la visualización de datos, revisa los fundamentos teóricos y analiza trabajos previos específicos en criminalidad, clasificando las principales herramientas y comparativas existentes para situar esta aportación en el marco de la literatura.

**Capítulo 3. Objetivos concretos y metodología de trabajo.** Define el objetivo general y los objetivos específicos del proyecto, y describe el enfoque metodológico seguido para la recogida, limpieza, análisis y evaluación de los datos, asegurando la reproducibilidad y el rigor del proceso.

**Capítulo 4. Marco normativo.** Detalla la normativa de protección de datos (GDPR) y las políticas de derechos de autor y reutilización de Eurostat, justificando la legalidad y la ética en el uso de los datos como base del trabajo.

**Capítulo 5. Desarrollo específico de la contribución.** Expone paso a paso la implementación práctica en Power BI, Tableau y Python (Plotly/Matplotlib), desde la preparación del entorno y la carga de datos hasta la generación de las visualizaciones y la evaluación técnica y funcional de cada herramienta.

**Capítulo 6. Código fuente y datos analizados.** Informa sobre la ubicación en GitHub del repositorio con todo el código desarrollado y los scripts de procesamiento, así como la inclusión de los datos de Eurostat en los ficheros de limpieza y generación de gráficos para garantizar la trazabilidad.

**Capítulo 7. Conclusiones.** Resume los hallazgos principales, relaciona los resultados obtenidos con los objetivos planteados y valora en qué medida se han cumplido, destacando la relevancia de la comparativa para la toma de decisiones en análisis de criminalidad.

**Capítulo 8. Limitaciones y prospectiva.** Reflexiona sobre las carencias detectadas durante el estudio y propone líneas de trabajo futuro para ampliar y profundizar la aportación.

## 2. Contexto y estado del arte

### 2.1.Contexto del problema

El análisis de la criminalidad en Europa es esencial para comprender las dinámicas delictivas y desarrollar estrategias efectivas de prevención y control. Según datos de Eurostat (2024), las estadísticas de delitos registrados por la policía varían significativamente entre los países europeos, lo que refleja diferencias en las definiciones legales, prácticas de registro y realidades sociales.

La visualización de datos desempeña un papel esencial en el análisis de la criminalidad, permitiendo identificar patrones y tendencias que facilitan la toma de decisiones informadas. Los analistas de delitos emplean métodos de representación cartográfica y análisis, como el análisis de puntos calientes, para identificar tendencias y pautas delictivas, lo que ayuda a los organismos policiales a detectar problemas, asignar recursos y resolver delitos. ArcGIS, por ejemplo, es una tecnología estándar en el sector para el análisis de la delincuencia y es fundamental para aplicar estrategias de reducción de la criminalidad basadas en datos y evidencias (Esri, s.f.).

**Figura 1.** Visualización de datos de delincuencia con ArcGIS.



Fuente: Esri, (s.f.).

La figura 1 muestra un mapa interactivo generado con **ArcGIS**, en el que se representan diferentes tipos de delitos geocalizados en una ciudad. Cada icono identifica un tipo de incidente (por ejemplo, robos, asaltos o vandalismo), permitiendo detectar patrones y zonas de concentración delictiva mediante una visualización clara y detallada sobre un entorno urbano.

Además, la aplicación de técnicas de análisis inteligente de datos y visualización ha demostrado ser efectiva en la investigación criminal, permitiendo identificar relaciones y patrones que podrían pasar desapercibidos en análisis tradicionales (Ribas et al., 2020).

En resumen, el análisis y visualización de datos criminalísticos son herramientas fundamentales para comprender y abordar la criminalidad en Europa, facilitando la implementación de políticas y estrategias basadas en evidencia.

## 2.2.Estado del arte

Hoy en día, con la cantidad de datos que se manejan en cualquier ámbito —desde empresas hasta instituciones públicas—, no basta con tener la información: hay que saber presentarla. La visualización de datos se ha convertido en una herramienta clave para entender mejor esa información y tomar decisiones más acertadas (Kelleher & Wagener, 2011). Y es que, cuando los datos se muestran de forma visual, clara y ordenada, todo cambia: se detectan patrones, se destacan problemas y se facilita mucho el análisis, incluso para personas que no son expertas en datos.

En este apartado se va a hacer un repaso general por las herramientas de visualización más conocidas y utilizadas actualmente, centrándose este estudio en tres: Tableau, Power BI y Python con Matplotlib. Se observará en qué se parecen, en qué se diferencian, cómo han evolucionado, y en qué contextos se están utilizando —incluido el análisis de criminalidad, que es el enfoque de este trabajo—. También se revisarán estudios y comparativas previas, para ver qué se ha hecho ya y qué huecos existen que justifiquen esta investigación.

### 2.2.1. Fundamentos de la visualización.

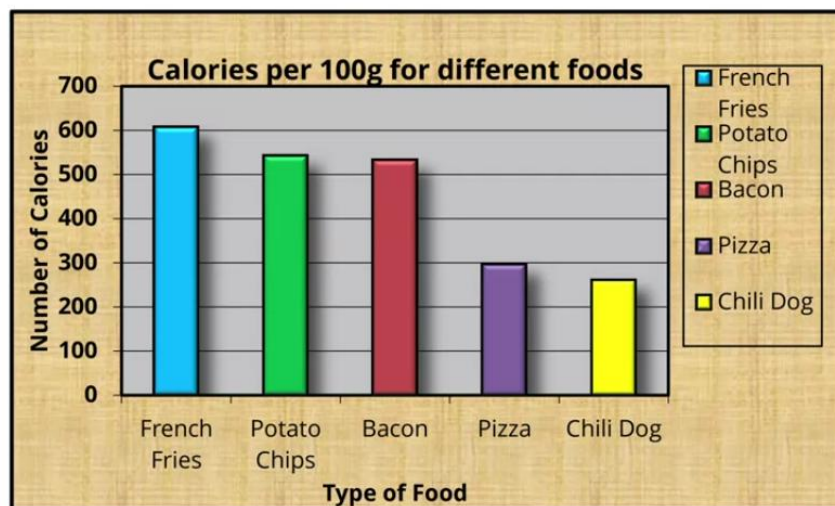
La visualización de datos tiene una larga trayectoria que comienza con los primeros mapas de la antigüedad, como el de Nippur, y continúa con figuras como Ptolomeo, Mercator o

Leonardo da Vinci, cuyo trabajo ya reflejaba el uso de representaciones gráficas para comprender el mundo (Gómez, 2002). En los siglos XVIII y XIX, la disciplina dio un gran impulso gracias a autores como Playfair, inventor de los gráficos de barras y líneas, y Minard, autor del célebre gráfico sobre la campaña de Napoleón que Tufte (2001, p. 40) consideró el mejor gráfico estadístico jamás creado. Desde entonces, la visualización ha evolucionado hasta convertirse en una herramienta esencial, especialmente en la era digital y del Big Data.

Se habla de los principios teóricos más aceptados hoy en día para crear visualizaciones útiles y efectivas, no se puede dejar de mencionar a Edward R. Tufte, uno de los autores más

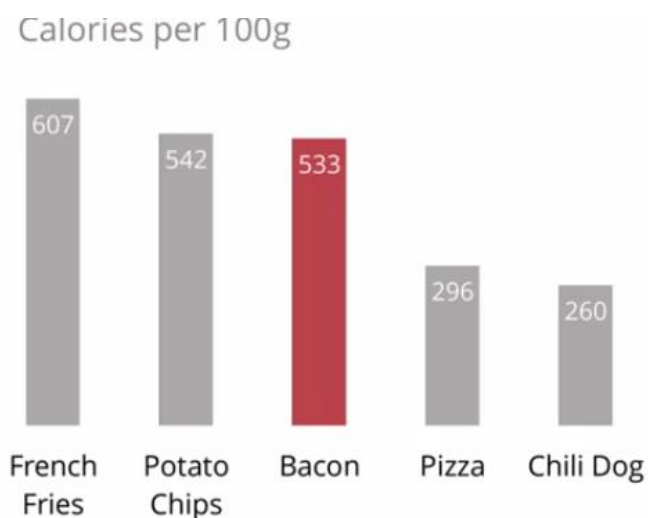
**Figura 2.** Ejemplo de gráfico con exceso de elementos decorativos.

influyentes en este campo. En *The Visual Display of Quantitative Information*, Tufte (2001) se defiende que las visualizaciones deben ser claras, precisas y evitar todo elemento innecesario que distraiga. Uno de sus aportes más conocidos es el concepto de “data-ink ratio”, que propone maximizar la cantidad de tinta (o elementos visuales) dedicada a representar datos reales, eliminando adornos superfluos. Según Tufte, una buena visualización debe centrarse en lo esencial y ser fiel a los datos, sin distorsiones ni exageraciones visuales.



Fuente: Raj, A. (2022)

**Figura 3.** Ejemplo de gráfico con la proporción de tinta ideal.



Fuente: Raj,A (2022).

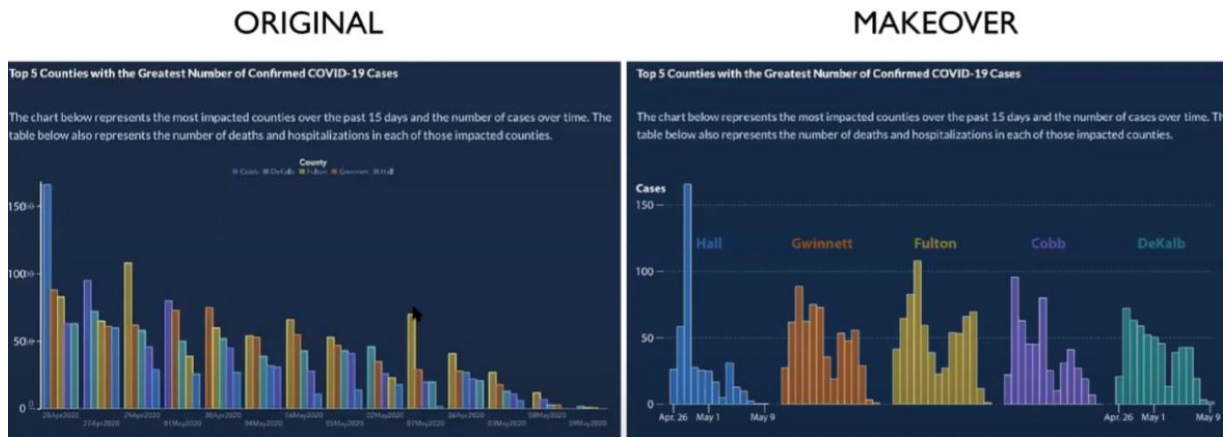
En las figuras 2 y 3 se muestra la comparativa entre dos gráficos que representan el mismo conjunto de datos (calorías por 100g en distintos alimentos), pero con enfoques visuales muy distintos. La figura 2 presenta un gráfico sobrecargado de elementos decorativos —como colores intensos, tramas de fondo y múltiples líneas— que distraen la atención y dificultan la lectura, lo que lo convierte en un mal ejemplo desde el punto de vista del diseño informativo. En cambio, la figura 3 simplifica la presentación visual eliminando lo superfluo y centrándose únicamente en los datos, siguiendo el principio del “data-ink ratio” propuesto por Tufte, lo que mejora considerablemente la comprensión y claridad de la información.

Otra figura que también cree que una buena visualización debe centrarse en la claridad y la utilidad es Stephen Few, uno de los principales defensores de un enfoque práctico y orientado al análisis. En *Now You See It*, Few (2009) insiste en que los gráficos deben diseñarse teniendo en cuenta cómo percibimos visualmente la información, priorizando la comprensión por encima del impacto visual. Para él, lo importante no es impresionar, sino ayudar a ver patrones, detectar anomalías y tomar decisiones basadas en datos. Por eso apuesta por gráficos simples, pero bien contruidos, como líneas, barras o tablas, y alerta sobre el riesgo de utilizar visualizaciones recargadas o poco intuitivas que, en lugar de aclarar, terminan confundiendo al usuario.

En esta misma línea, pero con un enfoque más narrativo y ligado al periodismo visual, destaca Alberto Cairo, uno de los grandes divulgadores de la visualización de datos en español. En *The*

*Functional Art*, Cairo (2012) defiende que una buena visualización no solo debe ser clara y precisa, sino también comprensible, estética y útil para contar historias. Plantea que el diseño visual tiene que equilibrar forma y función: debe ser visualmente atractivo, pero sin sacrificar la fidelidad a los datos ni dificultar la interpretación

**Figura 4.** Rediseño de un gráfico original del Departamento de Salud de Georgia sobre casos de COVID-19.



Fuente: Cairo, A (2020)

Como muestra Cairo (2021) en uno de sus rediseños, una buena organización de los datos y una jerarquía visual clara pueden transformar una visualización confusa en un gráfico mucho más comprensible y útil para el análisis. La imagen (figura 4) compara el gráfico original del Departamento de Salud de Georgia con su rediseño realizado por Alberto Cairo. En el diseño original, los datos están mezclados por fecha y condado, dificultando la lectura. En el makeover, Cairo agrupa la información por condado y la ordena cronológicamente, lo que permite interpretar mejor la evolución de los casos de COVID-19 y mejora considerablemente la claridad visual del conjunto.

Además, pone el foco en la responsabilidad ética del creador de visualizaciones, recordando que los gráficos también pueden manipular si no se diseñan con honestidad.

Con todo esto, queda claro que una visualización bien diseñada no es solo una cuestión estética, sino algo fundamental para entender los datos. Por eso es importante tener estos fundamentos bien interiorizados, especialmente cuando se trabaja con visualizaciones a mayor escala, ya que hoy en día hablar de datos sin mencionar el Big Data es prácticamente imposible. Como señalan Chen, Mao y Liu (2014), la visualización es una pieza clave dentro del campo del Big Data, ya que permite transformar datos complejos en algo comprensible y,



sobre todo, útil para los usuarios. Sin una buena visualización, el análisis de Big Data pierde gran parte de su valor.

#### 2.2.1.1. Panorama actual nacional: el impulso a la digitalización con el Kit Digital.

En un momento en el que visualizar claramente datos y el análisis digital se han convertido en aspectos clave para la toma de decisiones, la digitalización ya no es una opción, sino una necesidad, también para las pequeñas empresas. En España, este proceso ha ido cobrando especial relevancia en los últimos años gracias a iniciativas públicas como el Kit Digital, un programa impulsado por el Gobierno para facilitar el acceso de pymes y autónomos a soluciones tecnológicas. Este plan, enmarcado dentro del Plan de Recuperación, Transformación y Resiliencia y financiado por los fondos europeos NextGenerationEU, busca promover el desarrollo digital de las empresas españolas, con un presupuesto total de 3.067 millones de euros (AceleraPyme, s.f.).

El Kit Digital ha tenido un impacto significativo en la digitalización de las pymes y autónomos en España. Según datos del Plan de Recuperación, Transformación y Resiliencia, hasta octubre de 2024, cerca de 530.000 pymes han recibido ayudas por un total de 2.300 millones de euros a través de este programa. Con todo esto, se ha alcanzado el 80% del objetivo comprometido con Bruselas. Estas cifras reflejan la amplia acogida y efectividad del programa en el tejido empresarial español (2024).

El impacto de esta propuesta no solo se mide en número de beneficiarios, sino también en el tipo de soluciones tecnológicas que está promoviendo. Muchas de las categorías subvencionadas (como la gestión de procesos, el análisis de datos o los servicios de oficina virtual) están relacionadas con herramientas como Power BI, Tableau o soluciones en Python, lo que ha facilitado su adopción especialmente entre pequeñas empresas que antes no tenían acceso a este tipo de tecnologías. Gracias a estas ayudas, muchas pymes que antes no podían permitirse herramientas avanzadas de análisis y visualización han dado el salto a las plataformas analizadas en este trabajo. Esto ha contribuido a que estas tecnologías dejen de ser algo exclusivo de grandes corporaciones y pasen a formar parte del día a día de empresas más pequeñas. (AceleraPyme, s.f.).



### 2.2.2. Visualización en el análisis criminalístico.

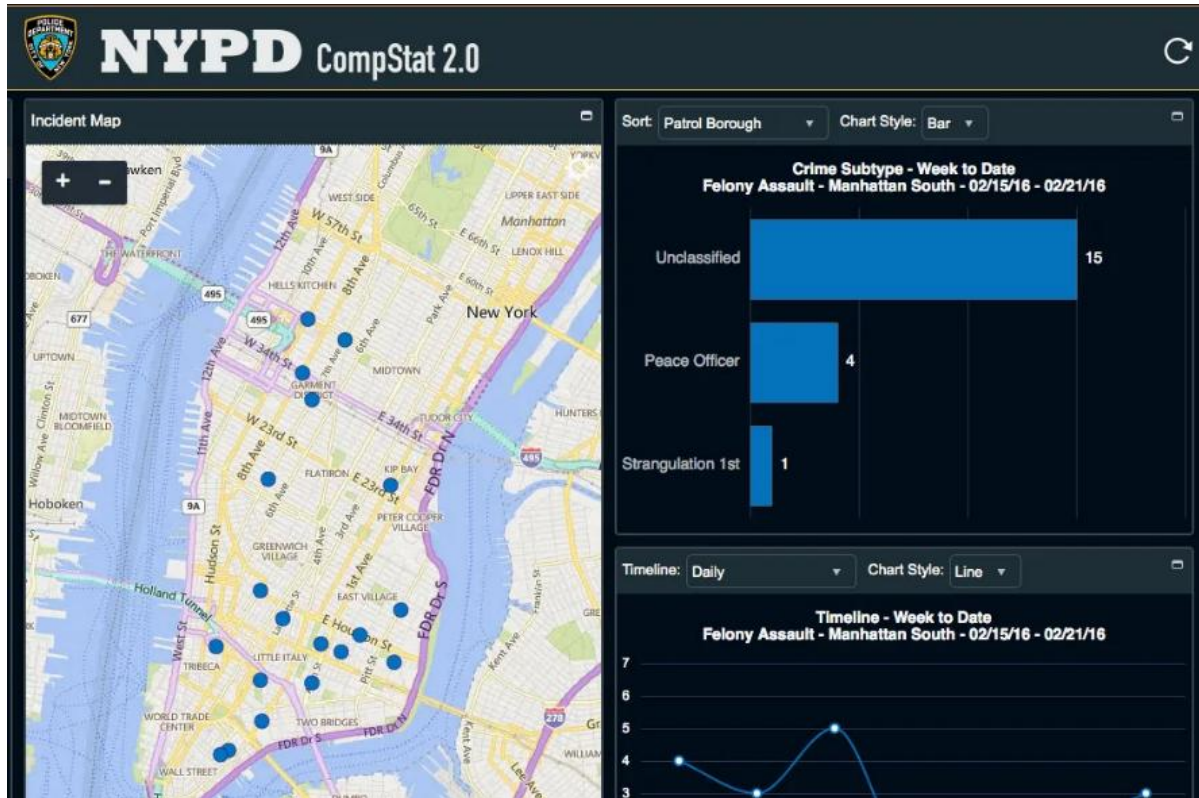
La visualización de datos también ha demostrado ser una herramienta clave en campos como la criminología, la seguridad ciudadana y el análisis policial. Gracias a su capacidad para mostrar patrones espaciales y temporales, se utiliza para identificar zonas de riesgo, analizar la evolución de distintos tipos de delitos o planificar intervenciones más eficaces por parte de las fuerzas de seguridad. Desde mapas de calor que muestran la concentración de robos en una ciudad, hasta dashboards que monitorizan en tiempo real la actividad delictiva, estas representaciones permiten tomar decisiones mejor informadas y, en muchos casos, anticiparse a los problemas.

En Europa se han desarrollado iniciativas relevantes que utilizan la visualización de datos para analizar fenómenos delictivos. El portal de Eurostat (s.f), por ejemplo, ofrece mapas y paneles interactivos que permiten explorar la evolución de distintos tipos de delitos registrados por la policía en los Estados miembros. Esto facilita la comparación entre países, ayudan a detectar zonas de mayor incidencia delictiva y permiten crear o modificar políticas públicas relacionadas con la seguridad ciudadana.

Si se observa fuera de Europa se pueden encontrar iniciativas destacadas como el sistema CompStat 2.0 del Departamento de Policía de Nueva York (NYPD). Esta plataforma pública ofrece estadísticas delictivas actualizadas semanalmente, presentadas mediante gráficos, tablas y mapas interactivos. Su objetivo es facilitar el análisis de los patrones delictivos por distrito, mejorar la transparencia institucional y reforzar la implicación ciudadana en la seguridad pública (NYPD, s.f.). Incluimos una imagen de la interfaz de CompStat 2.0 obtenida de fuentes abiertas en Internet, ya que la plataforma no permite el acceso desde determinadas ubicaciones fuera de Estados Unidos, lo que ha impedido su consulta directa desde Europa.

La figura 5 muestra la interfaz de CompStat 2.0, el sistema del Departamento de Policía de Nueva York que combina mapas interactivos con gráficos estadísticos. Permite visualizar la localización de delitos sobre el mapa y consultar, en paralelo, gráficos con la evolución temporal y la tipología de los crímenes. Esta herramienta facilita tanto el análisis interno como el acceso público a la información actualizada sobre la criminalidad en la ciudad.

**Figura 5.** Interfaz de CompStat 2.0 del Departamento de Policía de Nueva York, que permite a los ciudadanos acceder a estadísticas delictivas actualizadas y mapas interactivos.



Fuente: NYPD (s.f.)

Otro proyecto innovador en el campo de la visualización delictiva es CrimeRadar, desarrollado por el Instituto Igarapé en Brasil. Esta plataforma utiliza *machine learning* para predecir el riesgo de criminalidad en distintos barrios de Río de Janeiro, en base a datos recopilados por la policía durante más de cinco años. Fue lanzada durante los Juegos Olímpicos de 2016 y permite identificar zonas con mayor probabilidad de delitos en la semana siguiente (Igarapé Institute, s.f.).

La figura 6 corresponde a una visualización generada por la plataforma, que muestra en tiempo real el nivel de riesgo de criminalidad en distintas zonas de Río de Janeiro. Cada área se codifica por colores según su nivel de riesgo estimado, permitiendo identificar de forma visual cuáles son los sectores más seguros o peligrosos del entorno urbano.

**Figura 6.** Visualización del riesgo de criminalidad en tiempo real en un barrio de Río de Janeiro mediante la plataforma CrimeRadar.



Fuente: Instituto Igarapé (s.f.)

Además de los ejemplos relevantes ya mencionados, también existe una literatura aplicada que respalda el uso de la visualización de datos en la criminología. Aunque no es extensa, sí encontramos iniciativas institucionales destacadas. Por ejemplo, la *Agencia de los Derechos Fundamentales de la Unión Europea (FRA)* ha creado herramientas interactivas para explorar datos sobre victimización, violencia o percepción de inseguridad en distintos colectivos. De forma similar, la *Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC)* publica visualizaciones comparativas sobre delitos violentos, homicidios o tráfico de personas a nivel global. Este tipo de recursos refuerza el papel de la visualización como herramienta clave para comprender fenómenos criminales y orientar políticas públicas basadas en evidencia.

### 2.2.3. Herramientas de visualización: antecedentes y clasificación.

Las herramientas de visualización han crecido y se han diversificado en los últimos años. Para entender mejor el papel de Tableau, Power BI o Python, conviene clasificarlas según criterios como su modelo (comercial o abierto), su nivel de interactividad o el tipo de usuario al que se dirigen.

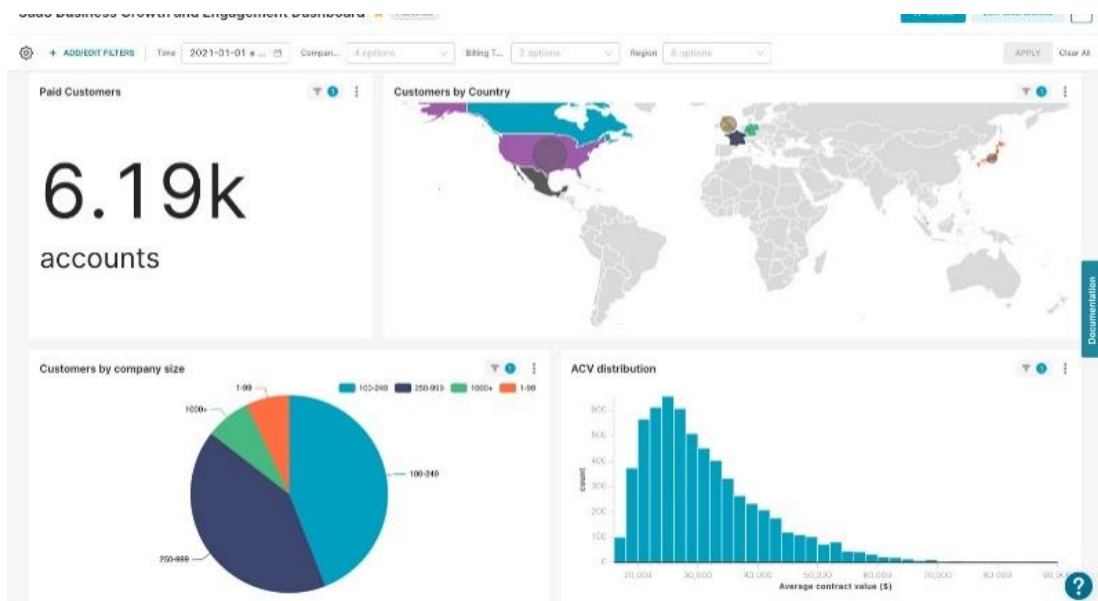
Una forma común de clasificar las herramientas de visualización es según su modelo de distribución. Las herramientas de pago, como Tableau, Power BI o Qlik Sense, suelen estar

orientadas a entornos empresariales y ofrecen interfaces visuales intuitivas, soporte técnico y funcionalidades avanzadas bajo licencia de pago.

Por otro lado, las herramientas de código abierto basadas en lenguaje de programación (y sus correspondientes librerías) como Python (con Matplotlib, Seaborn o Plotly), R (con ggplot2, Shiny) o Apache Superset, permiten un mayor control y personalización, aunque requieren conocimientos técnicos por parte del usuario (Tableau, s.f.; Microsoft Power BI, s.f.; Plotly, s.f.; Superset, s.f.).

La figura 7 muestra un dashboard interactivo generado con **Apache Superset**, una herramienta de código abierto para la exploración y visualización de datos. El panel incluye diferentes tipos de gráficos (mapas, histogramas y gráficos circulares) que permiten analizar métricas de negocio como número de clientes, distribución geográfica o tamaño de empresa, todo ello en una única vista integrada y personalizable.

**Figura 7.** Ejemplo de dashboard interactivo creado con Apache Superset



Fuente: Apache Superset (s.f)

Otro criterio importante la interactividad. Algunas herramientas ya mencionadas permiten crear, además de las estáticas, visualizaciones interactivas, en las que el usuario puede filtrar, explorar y modificar la vista de los datos, como ocurre con Tableau, Power BI o Plotly. En cambio, otras como Matplotlib o ggplot2 generan en su mayoría gráficos estáticos, ideales

para informes, publicaciones o cuando no se necesita que el usuario interactúe con las gráficas.

Para esta comparativa se han elegido Tableau, Power BI y Python con Matplotlib porque son tres herramientas muy representativas en el mundo de la visualización de datos. Tableau y Power BI destacan porque son visuales y fáciles de usar, ideales para aquellos usuarios que no tienen conocimientos técnicos, pero quieren crear dashboards interactivos de forma rápida. Tableau es conocida por el manejo visualizaciones complejas y lo flexible que es (Tableau, s.f.), mientras que Power BI encaja muy bien en entornos empresariales gracias a su integración con Excel y otros productos de Microsoft (Microsoft Power BI, s.f.).

Por otro lado, Python con Matplotlib representa la opción basada en código. Es más técnica pero también mucho más personalizable. Se usa muchísimo en ciencia de datos y en entornos académicos, y permite tener un control total sobre los gráficos (Hunter, 2007). Al combinar estas tres herramientas se pueden comparar soluciones visuales comerciales con una alternativa de código abierto, lo que permite ver qué ofrece cada una según el perfil del usuario y el tipo de proyecto.

Una vez contextualizado el panorama general de herramientas de visualización y justificada la elección de Tableau, Power BI y Python con Matplotlib, pasamos ahora a describir cada una de ellas con algo más de detalle. En los siguientes subapartados se analiza brevemente el origen, las principales características, ventajas y limitaciones de cada herramienta, con el objetivo de entender mejor su funcionamiento y el tipo de usuario o proyecto al que están más orientadas.

#### 2.2.3.1. Tableau.

Tableau fue desarrollado originalmente en 2003 por investigadores de la Universidad de Stanford y más adelante comercializado por la empresa Tableau Software. Su objetivo era facilitar la comprensión de los datos mediante una interfaz visual intuitiva, sin necesidad de programar. Desde entonces, ha evolucionado hasta convertirse en una de las herramientas de visualización más potentes y populares en el entorno empresarial. En 2019, fue adquirida por Salesforce, lo que reforzó su integración con plataformas de análisis en la nube (Tableau, s.f.).

Entre sus características más destacadas están la facilidad de uso mediante arrastrar y soltar (*drag and drop*), la amplia variedad de tipos de gráficos, la posibilidad de crear dashboards

interactivos y la conexión con múltiples fuentes de datos. Tableau es especialmente útil para usuarios no técnicos que necesitan explorar y comunicar datos de forma rápida y visual, aunque sus versiones más avanzadas también ofrecen funciones analíticas complejas.

Aunque Tableau está más extendido en entornos empresariales, también se ha utilizado en proyectos relacionados con la analítica criminal y la justicia social. Un ejemplo destacado es el proyecto Police Scorecard, que evalúa el rendimiento de departamentos de policía en EE. UU. en aspectos como uso de la fuerza, transparencia o desigualdad racial. El equipo del proyecto utilizó Tableau para representar grandes volúmenes de datos de forma clara y accesible, permitiendo detectar patrones preocupantes y proponer mejoras en las políticas de seguridad pública. En el caso de Police Scorecard, Tableau y la visualización de datos ofrecieron una solución potente para obtener información valiosa a partir de conjuntos de datos tan masivos. (Anzilotti, 2021)

Además, también destaca el caso de Andrew Wheeler, analista especializado en datos criminalísticos. En su portfolio se pueden encontrar numerosos dashboards interactivos y tutoriales centrados en el análisis de delitos, tanto con Excel como con Tableau. En 2021, desarrolló varios paneles de análisis criminal diseñados en Tableau, donde representaba patrones delictivos semanales y estacionales, con el objetivo de facilitar la toma de decisiones en el ámbito policial local (Wheeler, 2021).

Como se ha mencionado anteriormente, Tableau destaca por permitir la creación de visualizaciones interactivas sin necesidad de programar. El propio Andrew Wheeler (2021) señala que “la verdadera ventaja de Tableau frente a los gráficos estáticos de Excel (o cualquier otro programa estadístico) es que permite aplicar filtros interactivos y vincular visualizaciones entre sí”. No obstante, la herramienta también presenta algunas limitaciones. Por ejemplo, Tableau no ofrece la opción de actualizar automáticamente los informes, lo que dificulta su uso en entornos donde se requiere una programación regular. Además, los usuarios no pueden crear diseños estáticos o tablas de datos tradicionales, lo cual puede ser un inconveniente en ciertos contextos. Por último, la gran cantidad de funciones disponibles, aunque poderosa, puede resultar abrumadora para los usuarios principiantes (Thinklytics, s.f.). Es por ello que se debe considerar si es la mejor herramienta para usar en ciertos casos.

#### 2.2.3.2. Power BI.

Power BI es una herramienta de visualización y análisis de datos desarrollada por Microsoft. Surgió en 2013 como un conjunto de complementos para Excel (Power Query, Power Pivot y Power View), y en 2015 se lanzó oficialmente como producto independiente dentro del ecosistema de Microsoft 365. Desde entonces, ha crecido rápidamente, integrándose con otros servicios de la compañía como Azure y SharePoint, y posicionándose como una de las soluciones de inteligencia empresarial más utilizadas en entornos corporativos (Microsoft Power BI, s.f.).

Aunque Power BI se utiliza mayoritariamente en el ámbito empresarial, también ha encontrado un lugar destacado en instituciones públicas, donde se emplea para transformar grandes volúmenes de datos en visualizaciones claras y accesibles. Organismos como ayuntamientos, universidades o administraciones gubernamentales lo utilizan para hacer seguimiento de proyectos, gestionar recursos o tomar decisiones basadas en datos reales. Según Kumar (2021), su integración con Excel y su bajo coste han contribuido a que sea una herramienta muy accesible en entornos institucionales.

Además, en sectores como el de la seguridad, Power BI también ofrece ventajas específicas. Por ejemplo, permite centralizar la gestión de casos, facilitando el seguimiento de investigaciones, evidencias o el estado de cada expediente. Asimismo, puede aplicarse al análisis del tráfico para detectar puntos conflictivos, zonas con alta siniestralidad o patrones de congestión, ayudando así a mejorar la seguridad vial y la eficiencia en la respuesta de los servicios de emergencia (Kumar, 2021).

Power BI tiene un sinnúmero de ventajas que darían para varias páginas, pero una de las más destacadas es su bajo coste, ya que ofrece una versión gratuita y otras opciones de pago muy asequibles. Esto lo convierte en una elección accesible tanto para usuarios individuales como para organizaciones de distintos tamaños (QMetrix, s.f.). No obstante, también presenta algunas limitaciones, como ya ocurría con Tableau. Por ejemplo, puede resultar abrumador para algunos usuarios debido a la gran cantidad de funcionalidades, especialmente al trabajar con fórmulas complejas en DAX, lo que representa una barrera para quienes no tienen experiencia previa. Otra limitación importante es su compatibilidad limitada, ya que Power BI Desktop no tiene versión para Mac ni para Linux. Aunque se puede acceder a través de la versión web, esta puede presentar un rendimiento más lento y funcionalidad reducida en comparación con la versión de escritorio para Windows (Pavlenko, 2022).

### 2.2.3.3. Python con Matplotlib.

Python es un lenguaje de programación de alto nivel creado por Guido van Rossum y lanzado por primera vez en 1991. Con el paso del tiempo, se ha consolidado como uno de los lenguajes más populares en ciencia de datos y análisis gracias a su sintaxis sencilla, su gran comunidad y su amplia variedad de librerías. Una de las más destacadas en el ámbito de la visualización es Matplotlib, desarrollada inicialmente por John D. Hunter en 2003 con el objetivo de replicar la funcionalidad gráfica de MATLAB en Python. Desde entonces, Matplotlib se ha convertido en una herramienta de referencia para generar gráficos estáticos, simples y personalizables (Hunter, 2007).

A diferencia de herramientas como Tableau o Power BI, que se basan en interfaces visuales y permiten construir gráficos arrastrando elementos, Matplotlib en Python sigue un enfoque programático. Esto significa que el usuario debe escribir código para definir cada aspecto de la visualización: tipo de gráfico, ejes, colores, etiquetas, estilos, etc. Este enfoque ofrece una gran flexibilidad y control, lo que resulta ideal para usuarios con conocimientos técnicos que buscan personalizar sus gráficos al detalle. Sin embargo, también implica una mayor curva de aprendizaje y tiempos de desarrollo más largos en comparación con las soluciones visuales. Según McKinney (2012), el uso de Python para visualización es especialmente útil en entornos científicos y académicos, donde se valora la reproducibilidad y la automatización del análisis.

### 2.2.4. Comparativas previas.

Aunque hay bastantes análisis sobre herramientas de visualización de datos, la mayoría son bastante generales y no están centrados en áreas como la criminalidad o el uso institucional. Aun así, algunos estudios comparativos sirven para tener una idea clara del panorama actual. El primero, elaborado por Kadam y Akhade (2021), se centra en una revisión teórica y funcional de herramientas como Tableau, Power BI y bibliotecas de Python (Matplotlib y Seaborn). Evalúan aspectos como facilidad de uso, flexibilidad, opciones de personalización o conectividad con datos. Concluyen que no hay una herramienta "mejor", sino que la elección debe depender de las necesidades específicas del usuario y del tipo de análisis que se quiera realizar.

El segundo estudio, de Parthe (2023), compara exclusivamente Power BI y Tableau, analizando precios, interfaz, capacidades gráficas, integración con otras plataformas y colaboración. El



autor destaca que Power BI es más accesible económicamente e ideal si ya se usa el ecosistema Microsoft, mientras que Tableau ofrece mayor capacidad de visualización y flexibilidad, aunque a un coste más elevado. En ambos casos, se propone que la decisión entre una y otra debe basarse en el contexto, el presupuesto y los objetivos del usuario.

Con este panorama, los estudios revisados dan una visión general bastante útil sobre las diferencias entre las herramientas analizadas, aunque también tienen algunas limitaciones importantes. Por ejemplo, no entran a valorar cómo se aplican estas herramientas al análisis de datos criminalísticos, que suelen ser complejos y delicados. Tampoco evalúan si realmente ayudan a interpretar y comunicar bien las tendencias delictivas, algo clave cuando se trabaja con visualizaciones. Además, no se plantean si estas herramientas pueden ser útiles para instituciones públicas u organismos que trabajan con este tipo de datos, un enfoque que es precisamente el que se pretende cubrir en este trabajo.

### 2.3.Conclusiones

Aunque se han publicado bastantes comparativas entre herramientas de visualización, la mayoría se quedan en aspectos técnicos muy generales. Apenas hay trabajos que se centren en cómo estas herramientas funcionan cuando se aplican a datos criminalísticos reales, ni qué valor pueden aportar en contextos como la toma de decisiones públicas o el análisis institucional. Ahí es precisamente donde encaja este trabajo: en cubrir ese hueco, explorando no solo qué herramienta es más potente o fácil de usar, sino también cuál comunica mejor los datos cuando lo que está en juego es entender fenómenos delictivos y ayudar a gestionarlos de forma más eficaz.

A diferencia de otros estudios con un enfoque más teórico o centrado en especificaciones técnicas, este trabajo apuesta por una metodología práctica y aplicada. Se analizan las herramientas Tableau, Power BI y Python (Matplotlib) a través de su uso real sobre un conjunto de datos relacionados con criminalidad en Europa. El objetivo no es solo evaluar las funciones que ofrecen, sino también observar cómo se comportan en la práctica: su facilidad de uso, la calidad de las visualizaciones generadas y su capacidad para transmitir información compleja de forma clara y comprensible. Esta aproximación permite ofrecer una comparativa más realista y útil, especialmente orientada a investigadores, analistas o instituciones públicas interesadas en utilizar estas herramientas para entender mejor los fenómenos delictivos.

## 3. Objetivos concretos y metodología de trabajo

### 3.1. Objetivo general

Comparar y aplicar tres herramientas de visualización de datos (Tableau, Power BI y Python con Plotly/Matplotlib) para el análisis de tendencias criminales en Europa, con el propósito de identificar cuál ofrece una mejor capacidad para representar y comunicar información compleja de forma clara, eficaz y accesible. Este trabajo analizará la usabilidad, la calidad visual y el rendimiento de cada solución, ofreciendo una guía práctica que facilite la toma de decisiones basada en datos en contextos institucionales o investigativos.

### 3.2. Objetivos específicos

Una vez definido el objetivo general del trabajo, este se desglosa en una serie de objetivos específicos que permitan abordar de forma estructurada las fases del análisis. A continuación, se enumeran los objetivos específicos planteados para este estudio:

- Identificar las características clave de cada herramienta de visualización (Tableau, Power BI y Python con Plotly/Matplotlib) en relación con su aplicabilidad al análisis de datos criminalísticos.
- Comparar la facilidad de uso, el rendimiento y la calidad de las visualizaciones ofrecidas por cada herramienta.
- Desarrollar visualizaciones representativas a partir de un conjunto de datos criminalísticos europeos.
- Evaluar la efectividad de cada herramienta en la interpretación y comunicación de tendencias delictivas.
- Valorar si cada herramienta puede ser útil para instituciones públicas o entidades que analizan la criminalidad.
- Resumir los resultados y proponer qué herramienta es más recomendable según el tipo de usuario o situación.

### 3.3. Metodología del trabajo

Para realizar el análisis y la comparación de las tres herramientas de visualización (Tableau, Power BI y Python con Matplotlib), se ha optado por CRISP-DM como marco general de trabajo. Esta metodología facilita la organización de las fases de recolección y limpieza de datos, la creación de visualizaciones y, finalmente, la evaluación de los resultados de forma ordenada y coherente (IBM, s.f.). Se le añadirá complementariamente un sistema de evaluación con enfoque propio, con multicriterio y especialmente diseñado para proyectos del tipo 4 como este. Mediante este enfoque, se asignan criterios relevantes para cada herramienta, asegurando la comparación ideal para los objetivos de este estudio. Según IBM (s.f.), las fases para llevar a cabo esta metodología son:

1. Entendimiento y preparación de los datos: Se identifican las fuentes de datos criminalísticos europeos y se aplican procesos de limpieza y transformación para garantizar su coherencia y calidad. (En CRISP-DM sería la fase Business + Data Understanding y Data Preparation).

2. Creación de visualizaciones: Se elaboran las representaciones gráficas con cada herramienta (Tableau, Power BI y Python con Plotly/Matplotlib), siguiendo recomendaciones de diseño que favorecen la claridad y la comprensión de la información. (En CRISP-DM la fase Modelling)

3. Comparación de resultados: Se evalúan algunos criterios, aplicando un sistema de evaluación multicriterio adaptado a los objetivos del estudio. (En CRISP-DM la fase Evaluation). Los criterios que podrían ser evaluados son:

- Facilidad de uso
- Calidad visual
- Rendimiento
- Flexibilidad de análisis
- Curva de aprendizaje, entre otros.

4. Conclusión y recomendación final: Se resumen los hallazgos y se determina qué herramienta ofrece mejores resultados dependiendo de los casos de uso, como orientación para futuros analistas. (En IBM la fase Deployment)

Con esta combinación entre CRISP-DM y el sistema de evaluación multicriterio, se busca la rigurosidad en las fases de análisis de datos con una valoración detallada de cada tecnología, de modo que resulte posible ofrecer una visión práctica y fundamentada sobre cuál de las soluciones se adapta mejor a distintos contextos de uso.

## 4. Marco normativo

En este trabajo se emplean datos procedentes de Eurostat, la oficina estadística de la Unión Europea, cuyos conjuntos de datos están diseñados para ser públicos, agregados y anonimizados, de modo que no contienen información que permita la identificación de personas físicas. Según la Guía de uso de Eurostat, todos los datos disponibles en las secciones “Tables” y “Database” cumplen los principios de licitud, transparencia y minimización del Reglamento UE 2016/679 (GDPR), al no incluir microdatos identificativos. Cuando está permitido el acceso a microdatos individuales, este se limita a usuarios registrados con fines científicos y bajo estrictas condiciones de anonimización, garantizando así el respeto a la normativa de protección de datos.

Eurostat fomenta la libre reutilización de sus datos estadísticos tanto para fines comerciales como no comerciales, sin necesidad de licencia previa, salvo en los casos en que los derechos de autor pertenezcan a terceros o se incluyan logotipos protegidos (Cartoy García, 2023). Esta política se basa en el aviso legal de la Comisión Europea y contribuye a que el uso de la información en este TFM sea totalmente lícito y transparente. Además, según Cartoy (2023), cada tabla indica la fecha de su última actualización y el periodo cubierto, lo que permite documentar la trazabilidad de las consultas y asegurar que los análisis se realizan sobre datos recientes.

Para la extracción y automatización de datos, Eurostat ofrece múltiples herramientas: descarga masiva en formato TSV, servicios web API basados en SDMX 2.1 y JSON, así como un navegador de datos interactivo integrado en la plataforma (Cartoy García, 2023). Estas opciones facilitan la adquisición reproducible de series temporales y garantizan que los procedimientos de obtención de datos cumplan con las mejores prácticas de calidad y seguridad, alineados con los estándares internacionales de estadística.

## 5. Desarrollo específico de la contribución

### 5.1. Planteamiento de la comparativa.

#### 5.1.1. Identificación del problema y motivación

Tal como se desarrolló en el apartado del contexto y estado del arte, la visualización de datos desempeña un papel esencial en el análisis de fenómenos extensos y complejos como la criminalidad, ya que permite representar patrones de forma más clara e intuitiva que el análisis puramente numérico. En este trabajo se ha optado por aplicar herramientas de visualización para facilitar el estudio comparativo de distintos indicadores relacionados con la criminalidad en Europa.

Para ello, se han seleccionado tres datasets del portal oficial de Eurostat: uno sobre asaltos sexuales registrados por la policía, otro sobre homicidios intencionados y un tercero que recoge el número de prisioneros por país y año. Todos los datasets contienen variables temporales y geográficas, así como datos de frecuencia. Sin embargo, en su formato original presentan muchas columnas irrelevantes para el análisis propuesto que serán depuradas más adelante durante la fase de limpieza. Además, con el objetivo de no sobrecargar los informes ni dificultar la lectura de las gráficas, se ha decidido limitar el análisis a seis países europeos seleccionados, que van a permitir un equilibrio visual en las comparativas.

#### 5.1.2. Soluciones seleccionadas y criterios de evaluación

Este trabajo se centra en la comparación de tres herramientas, tal y como se introdujo previamente, ampliamente utilizadas en el ámbito de visualización de datos: Tableau, PowerBI y Python con la biblioteca Matplotlib. Mientras que las dos primeras son plataformas comerciales diseñadas para usuarios sin conocimientos técnicos avanzados, con interfaces gráficas intuitivas, Python representa un enfoque más técnico basado en programación, que ofrece mayor control y personalización a costa de una mayor complejidad de uso. Esta diferencia de enfoque entre herramientas BI y herramientas de código permite explorar cómo se adaptan a distintos perfiles de usuario y a diferentes necesidades analíticas.

Para evaluar estas herramientas se ha seguido el modelo metodológico CRISP-DM, aplicando en su fase de evaluación una serie de criterios adaptados a los objetivos del estudio. Se ha considerado la facilidad de uso, entendida como la sencillez en la carga de datos, creación de

gráficos y navegación por el entorno; la calidad visual, basada en la estética y claridad de las visualizaciones generadas por defecto; el rendimiento, medido en función del tiempo de carga y la fluidez al trabajar con conjuntos de datos de tamaño medio; la flexibilidad de análisis, es decir, la capacidad de personalizar filtros, cálculos, formatos y tipos de gráficos; y, por último, la curva de aprendizaje, que hace referencia al esfuerzo requerido para que un usuario sin experiencia previa pueda desenvolverse de forma autónoma en la herramienta.

## 5.2.Desarrollo de la comparativa

### 5.2.1. Preparación del entorno y carga de datos

Se han seleccionado tres datasets relacionados con la criminalidad en Europa, descargados del portal oficial de Eurostat en formato CSV. Aunque estaban disponibles en otros formatos, el CSV resulta el más adecuado para su lectura y manipulación en las tres herramientas de análisis empleadas en este trabajo.

El primer dataset hace referencia a los asaltos sexuales registrados por la policía, y contiene variables como el país, el año, la frecuencia y el sexo de las víctimas. Además, incorpora una columna que especifica el tipo de valor representado (por ejemplo, número absoluto), cuyo dato numérico se encuentra en la columna OBS\_VALUE, que recoge el número total de asaltos sexuales por país y año. El archivo también incluye columnas adicionales de metadatos, como DATAFLOW, OBS\_FLAG o CONF\_STATUS, que no aportan información útil para el análisis y serán eliminadas en la fase de limpieza posterior.

**Figura 8.**Visión general del dataset de violencia sexual.

DATAFLOW	LAST UPDATE	freq	iccs	unit	geo	TIME_PERIOD	OBS_VALUE	OBS_FLAG	CONF_STATUS
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2014	97		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2015	149		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2016	120		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2017	113		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2018	109		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2019	154		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2020	145		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2021	135		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2022	169		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Albania	2023	134		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2014	3564		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2015	3479		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2016	4391		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2017	4253		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2018	4425		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2019	4509		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2020	4141		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2021	4354		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2022	4968		
STAT:CRIM_OFF_CAT(1.0)	23/04/2025 11:00:00	Annual	Sexual violence	Number	Austria	2023	5447		

Fuente: Elaboración propia.

En la imagen (figura 8) se muestra un extracto del conjunto de datos sobre asaltos sexuales, con registros organizados por país y año. Las columnas incluyen información estructural como la frecuencia temporal del dato (freq), la categoría delictiva (iccs), el tipo de unidad (unit), el país (geo), el año del registro (TIME\_PERIOD) y el valor observado (OBS\_VALUE).

El segundo dataset seleccionado recoge datos sobre **homicidios intencionados** en países europeos. Su estructura es muy similar al set de datos anterior, incluyendo variables como el país, el año y el sexo, junto con otras columnas de metadatos que no se describen de nuevo por ser equivalentes. Este archivo también presenta el número de homicidios registrados por país y año, lo que permite analizar su evolución temporal y compararla con otros indicadores relacionados, como la población penitenciaria. En la siguiente imagen (figura 9) se puede apreciar que la estructura del dataset es prácticamente la misma.

**Figura 9.** Visión general del dataset de homicidios.

DATAFLOW	LAST UPDATE	freq	iccs	unit	geo	TIME_PERIOD	OBS_VALUE	OBS_FLAG	CONF_STATUS
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2014	98		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2015	54		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2016	71		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2017	52		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2018	51		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2019	58		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2020	52		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2021	61		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2022	42		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania	2023	34		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2014	43		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2015	42		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2016	49		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2017	61		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2018	73		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2019	74		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2020	54		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2021	59		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2022	65		
ESTAT:CRIM_OFF_CATSDEFAULTVIEW(1.0)	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria	2023	75		

Fuente: Elaboración propia.

El tercer y último dataset seleccionado corresponde a la población penitenciaria en países europeos. Al igual que los anteriores, presenta una estructura similar con columnas que identifican el país, el año y el número total de personas en prisión. En este caso, el archivo ofrecía la posibilidad de desagregar los datos por sexo y grupo de edad; sin embargo, se ha optado por unificar todos los registros con el fin de simplificar la visualización y facilitar la

interpretación de los gráficos posteriores. A continuación, se muestra una imagen (Figura 10) con una vista preliminar de las primeras filas del dataset.

**Figura 10.** Visión general del dataset de prisioneros.

DATAFLOW	LAST UPDATE	freq	age	sex	unit	geo	TIME_PERIOD	OBS_VALUE	OBS_FLAG	CONF_STATUS
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2013	4998		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2014	5689		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2015	5981		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2016	6031		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2017	5674		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2018	5316		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2019	5045		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2020	4614		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2021	4956		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Albania	2022	5060		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2013	8862		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2014	8692		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2015	8665		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2016	8619		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2017	8852		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2018	9163		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2019	9072		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2020	8488		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2021	8436		
ESTAT:CRIM_PRIS_AGE(1.0)	23/04/2025 11:00:00	Annual	Total	Total	Number	Austria	2022	8993		

Fuente: Elaboración propia.

## POWERBI

La carga de los datos en Power BI se ha realizado a través de la funcionalidad “Obtener datos”, seleccionando la opción de archivo CSV para cada uno de los tres datasets. Una vez importados, Power BI permite realizar una vista previa de los datos y acceder a Power Query, herramienta fundamental para la limpieza y transformación previa al análisis.

**Figura 11.** Interfaz Power Query

= Table.TransformColumnTypes("#Encabezados promovidos",{{"DATAFLOW", type text}, {"LAST UPDATE", type datetime}, {"freq", type text},						
	DATAFLOW	LAST UPDATE	freq	iccs	unit	geo
1	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
2	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
3	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
4	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
5	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
6	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
7	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
8	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
9	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
10	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Albania
11	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria
12	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria
13	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria
14	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria
15	ESTAT:CRIM_OFF_CAT\$DEFAULTVIEW(...	23/04/2025 11:00:00	Annual	Intentional homicide	Number	Austria

Fuente: Elaboración propia.



En esta imagen (figura 11) se muestra la interfaz de Power Query tras la carga inicial del dataset. De forma automática, Power BI aplica ciertos ajustes básicos, como la detección del tipo de dato en cada columna y la promoción de la primera fila como encabezado. Estas acciones automáticas resultan especialmente útiles, ya que simplifican las tareas iniciales de limpieza y permiten trabajar con una estructura de datos más organizada desde el primer momento.

Como primer paso en la preparación de los datos, se ha realizado un proceso común en los tres datasets que incluye, además de la eliminación de columnas irrelevantes, la conversión del tipo de datos en columnas como LAST\_UPDATE y geo, asignándoles respectivamente los formatos de fecha y localización geográfica adecuados. Asimismo, se han renombrado varias columnas con el objetivo de facilitar su comprensión durante el análisis posterior.

En cuanto a los registros, se ha llevado a cabo una selección específica de países, descartando aquellos que no aportaban valor directo al estudio o cuya información era incompleta o inconsistente. Se han conservado los datos de seis países europeos con distintos perfiles sociales y criminales, que permiten realizar comparativas representativas:

- España, como país de referencia del estudio, para tener una visión clara de la situación nacional.
- Suecia, por ser un país con políticas sociales avanzadas y una percepción pública de alta seguridad.
- Países Bajos, por su enfoque progresista en justicia y su baja tasa histórica de encarcelamiento.
- Alemania, como una de las mayores potencias europeas, con un sistema judicial sólido y bien documentado.
- Rumanía, por representar contextos socioeconómicos distintos dentro de la Unión Europea y por su evolución reciente en justicia criminal.
- Lituania, debido a sus altas tasas de criminalidad registradas en algunos años y su interés desde una perspectiva comparativa.

Por último, en la variable temporal año, se ha acotado el análisis al periodo comprendido entre 2014 y 2022. Esta decisión responde a la necesidad de mantener coherencia temporal entre

los tres datasets y garantizar que los datos sean comparables en todas las herramientas utilizadas.

A continuación, se ha diseñado un modelo en estrella para estructurar los datos de forma más eficiente y facilitar la creación de visualizaciones complejas. Para ello, se han creado dos tablas de dimensiones: una de países y otra de años (mostradas en la figura 12 y 13), que se relacionan con las tablas de hechos correspondientes a cada uno de los datasets (asaltos sexuales, homicidios intencionados y prisioneros). Este modelo relacional permite unificar filtros y segmentaciones en los informes, asegurando la consistencia en el análisis cruzado de los diferentes indicadores.

**Figura 12.**Tabla dimension localización.

	A <sup>B</sup> <sub>C</sub> Localizacion	I <sup>2</sup> <sub>3</sub> id_localizacion
1	Germany	1
2	Spain	2
3	Lithuania	3
4	Netherlands	4
5	Romania	5
6	Sweden	6

Fuente: Elaboración propia.

**Figura 13.**Tabla dimensión año.

	I <sup>2</sup> <sub>3</sub> Año	I <sup>2</sup> <sub>3</sub> id_año
1	2014	1
2	2015	2
3	2016	3
4	2017	4
5	2018	5
6	2019	6
7	2020	7
8	2021	8
9	2022	9
10	2023	10

Fuente: Elaboración propia.

Como resultado, obtenemos estas tres tablas de hechos correctamente estructuradas (figura 14, 15 y 16) , con las que ya podremos trabajar en el apartado de visualizaciones de PowerBI:

**Figura 14.**Tabla hechos asaltos sexuales.

1 <sup>2</sup> <sub>3</sub> id_localizacion	1 <sup>2</sup> <sub>3</sub> id_año	1 <sup>2</sup> <sub>3</sub> Tasa_prisioneros
1	1	63228
2	1	65017
1	2	63020
2	2	61614
1	3	64291
1	4	65841
1	5	65762
1	6	63146
1	7	60505

Fuente: Elaboración propia.

**Figura 15.**Tabla hechos prisioneros.

AB <sub>C</sub> iccs	1 <sup>2</sup> <sub>3</sub> id_localizacion	1 <sup>2</sup> <sub>3</sub> id_año	1 <sup>2</sup> <sub>3</sub> Tasa_violencia_sexual
Sexual violence	1	1	34959
Sexual violence	2	1	9468
Sexual violence	1	2	34265
Sexual violence	1	3	37166
Sexual violence	1	4	34815
Sexual violence	1	5	40585
Sexual violence	1	6	40724
Sexual violence	1	7	41079
Sexual violence	1	8	41790
Sexual violence	1	9	49690

Fuente: Elaboración propia.

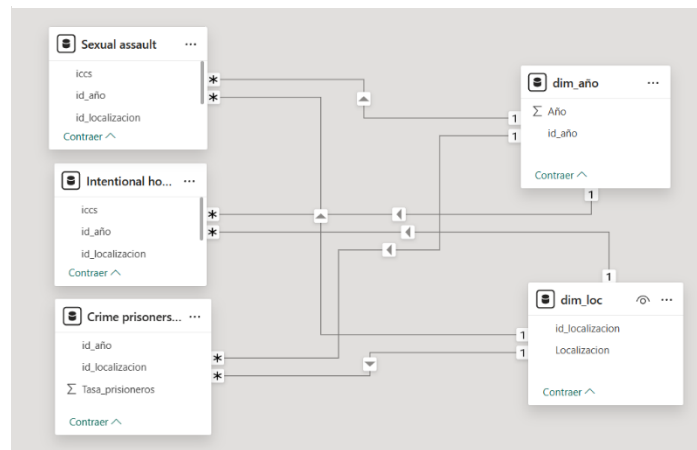
**Figura 16.**Tabla hechos homicidios.

AB <sub>C</sub> iccs	1 <sup>2</sup> <sub>3</sub> id_localizacion	1 <sup>2</sup> <sub>3</sub> id_año	1 <sup>2</sup> <sub>3</sub> Tasa_homicidio
Intentional homicide	1	1	645
Intentional homicide	2	1	323
Intentional homicide	1	2	655
Intentional homicide	1	3	747
Intentional homicide	1	4	738
Intentional homicide	1	5	632
Intentional homicide	1	6	586
Intentional homicide	1	7	719
Intentional homicide	1	8	631

Fuente: Elaboración propia.

Antes de comenzar con la fase de visualización, es fundamental comprobar que el modelo en estrella se encuentra correctamente relacionado. Una configuración incorrecta en las relaciones entre tablas puede generar inconsistencias en los resultados, afectar a la integridad de los filtros y, en consecuencia, producir visualizaciones erróneas o difíciles de interpretar.

**Figura 17.**Vista del modelo estrella



Fuente: Elaboración propia.

Por ello, se ha verificado que todas las relaciones entre las tablas de hechos y las dimensiones de país y año estén correctamente definidas, garantizando así la fiabilidad del análisis posterior. En la figura 17 se puede ver que las tres tablas de hechos —Sexual assault, Intentional homicides y Crime prisoners— se conectan con las dos dimensiones comunes: dim\_año y dim\_loc. Las relaciones están correctamente configuradas con claves primarias en las dimensiones (id\_año, id\_localizacion) y claves externas en las tablas de hechos, lo que garantiza la integridad del modelo y la coherencia en el análisis cruzado de los datos.

### PYTHON.

En Python, el proceso de carga y limpieza de los datos se ha realizado utilizando la biblioteca Pandas, que permite trabajar con estructuras de datos tipo *dataframe* de manera eficiente. Cada uno de los tres archivos CSV ha sido importado mediante la función `read_csv()`, y posteriormente se han aplicado los mismos pasos de limpieza descritos en Power BI: eliminación de columnas irrelevantes como `DATAFLOW`, `OBS_FLAG`, `CONF_STATUS` o `LAST_UPDATE`, conversión de tipos de datos en las columnas clave (como `TIME_PERIOD` a tipo fecha y geo como texto), y filtrado de registros para quedarnos solo con los seis países seleccionados.

A diferencia de Power BI, en este entorno no se ha creado un modelo en estrella con tablas de dimensiones, ya que en Pandas se trabaja principalmente con estructuras de datos planas. Sin embargo, se han renombrado las columnas principales para facilitar el análisis posterior y mantener la coherencia con el resto de herramientas utilizadas.

**Figura 18.** Carga y limpieza con Python.

```
[1]: import pandas as pd
df_prisioneros= pd.read_csv('Crime prisoners by age and sex.csv')
df_homicidios= pd.read_csv('Intentional homicide per country.csv')
df_violenciasex= pd.read_csv('Sexual assault.csv')
```

```
[13]: #Limpiamos columnas del df de prisioneros.
columnas_eliminadas= ['DATAFLOW', 'LAST UPDATE', 'freq', 'age', 'sex', 'unit', 'OBS_FLAG', 'CONF_STATUS']
df_prisioneros= df_prisioneros.drop(columns= columnas_eliminadas)
df_prisioneros.head()
```

```
[13]:
```

	geo	TIME_PERIOD	OBS_VALUE
0	Albania	2013	4998.0
1	Albania	2014	5689.0
2	Albania	2015	5981.0
3	Albania	2016	6031.0
4	Albania	2017	5674.0

```
[23]: #Quitamos paises que no quiero
países_deseados= ['Sweden', 'Netherlands', 'Romania', 'Spain', 'Lithuania', 'Germany']
df_prisioneros = df_prisioneros[df_prisioneros['geo'].isin(países_deseados)]
#df_prisioneros['geo'].unique()
df_prisioneros.head()
```

Fuente: Elaboración propia.

En la Figura 18 se muestra un fragmento del proceso de carga y limpieza del dataset de prisioneros utilizando la biblioteca Pandas. En este ejemplo se ilustran las operaciones de importación mediante `read_csv()`, así como la eliminación de columnas y registros que no son relevantes para el análisis. No se repiten los pasos realizados con los otros dos datasets, ya que el procedimiento ha sido análogo en todos los casos.

Dado que en Python no se pueden definir relaciones entre tablas como en los modelos estrella de Power BI, se ha optado por realizar un merge entre los distintos *dataframes*, unificando los datos en un único conjunto. Esto permite trabajar de manera más sencilla en la fase de visualización. El resultado final de esta fusión puede observarse en la Figura 19, donde se aprecia la estructura consolidada del dataset combinado.

**Figura 19.**Dataset final en Python

```
#Juntamos los 3 datasets limpios
#primero junto el df_prisioneros y el de df_violenciasex
df_combinado = df_prisioneros.merge(df_violenciasex, on=['País', 'Año'])

#Merge con homicidios
df_completo = df_combinado.merge(df_homicidios, on=['País', 'Año'])

df_completo.head()
```

	País	Año	Tasa_prisioneros	Tasa_asaltos_sexuales	Tasa_homicidios
0	Germany	2014	63228.0	34959	645.00
1	Germany	2014	63228.0	34959	0.80
2	Germany	2014	63228.0	34959	34959.00
3	Germany	2014	63228.0	34959	43.28
4	Germany	2014	63228.0	34959	7345.00

Fuente: Elaboración propia.

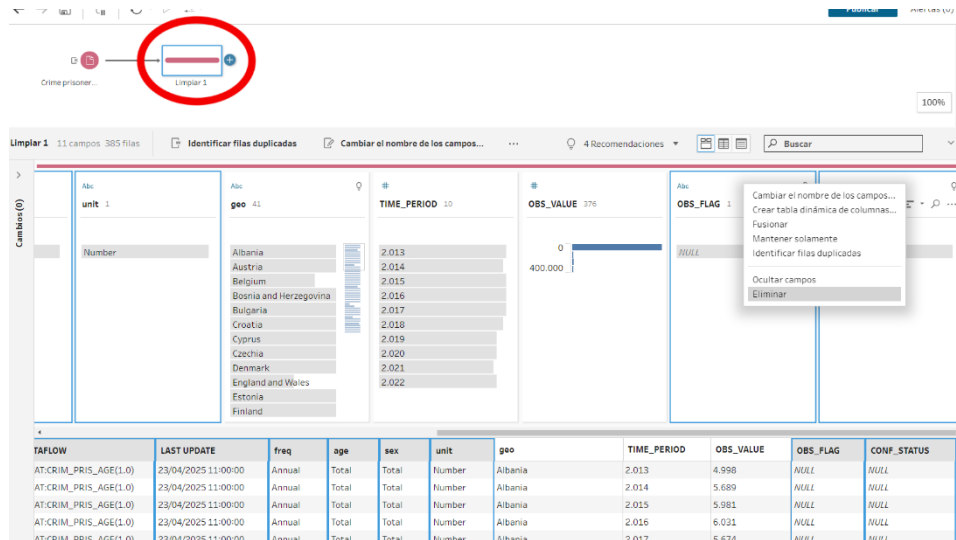
## TABLEAU

Dado que el propio Tableau no dispone de herramientas de transformación y limpieza de datos tan potentes como Power BI o las bibliotecas de Python, se ha optado por utilizar Tableau Prep, la herramienta complementaria de Tableau diseñada específicamente para la preparación de datos. Tableau Prep permite construir flujos visuales de transformación mediante una interfaz gráfica en la que cada paso del proceso (como limpieza, unión, filtrado o agregación) se representa de forma clara y modular.

A diferencia de Power BI, donde los pasos se aplican de forma más directa y secuencial, en Tableau Prep es necesario estructurar el flujo completo desde el inicio, añadiendo cada paso de forma explícita. Esto permite tener un mayor control sobre el proceso, pero también

implica una planificación más cuidadosa, ya que cualquier modificación requiere reorganizar o eliminar pasos concretos del flujo.

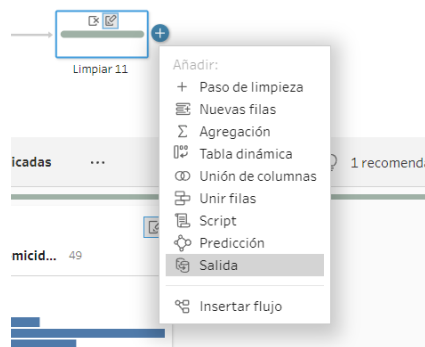
**Figura 20.** Paso de limpieza en Tableau.



Fuente: Elaboración propia.

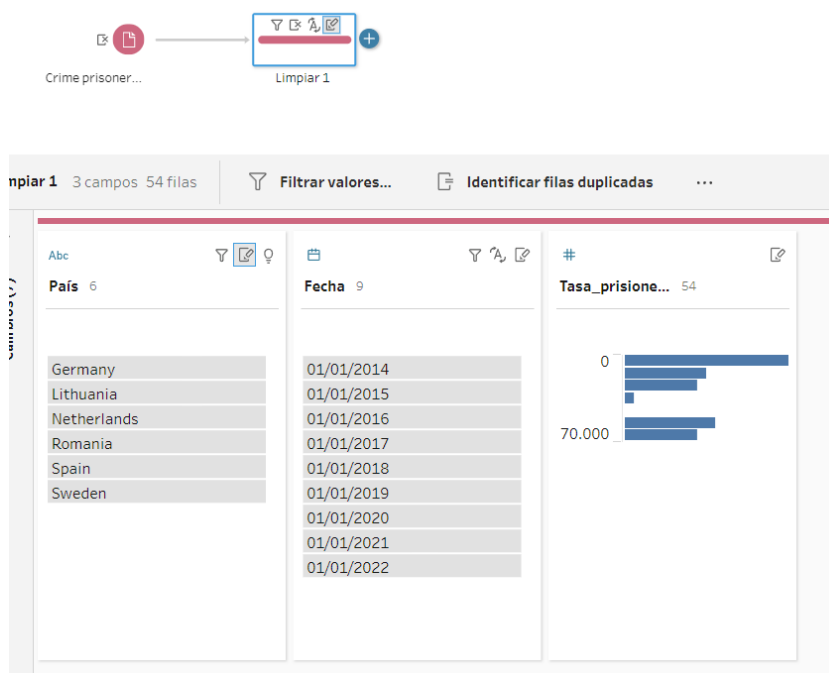
En la Figura 20 se muestra la interfaz de Tableau Prep, donde es posible visualizar el flujo de trabajo y los pasos aplicados a lo largo del proceso de limpieza. En la parte superior de la pantalla se encuentran las opciones para añadir nuevos pasos de transformación, como filtrado, agrupación o cambio de tipo de dato. Es importante destacar que, una vez finalizada la limpieza y transformación del dataset —replicando los mismos criterios aplicados en Power BI y Python—, se debe añadir un paso final de “salida”. Este paso es fundamental para exportar el resultado del flujo y poder utilizar los datos procesados posteriormente en Tableau Desktop para la elaboración de visualizaciones. En la figura 21 se muestra el paso fundamental de salida.

**Figura 21.** Paso de salida en Tableau



Fuente: Elaboración propia.

**Figura 22. Tabla de prisioneros limpia en Tableau.**



Fuente: Elaboración propia.

En la Figura 22 se muestra un ejemplo del resultado final de una tabla ya preparada en Tableau Prep, lista para ser utilizada en Tableau Desktop. Como puede observarse, todos los pasos de limpieza y transformación se han concentrado en un único bloque. Esta forma de trabajar puede presentar ciertas limitaciones, ya que en caso de necesitar eliminar o modificar alguna operación específica, habría que rehacer todo el paso completo. No obstante, dado que el proceso de depuración ha sido sencillo y replicado en todas las herramientas, y que los criterios aplicados eran claros y homogéneos, se ha considerado eficiente agrupar todas las operaciones en una única fase.

Cabe destacar que el tiempo de carga de los datos en Tableau Prep ha sido notablemente superior al observado en Python o Power BI, lo que ha supuesto una cierta dificultad adicional en la manipulación inicial de los datasets. A pesar de ello, una vez configurado el flujo, la herramienta permite exportar los datos limpios directamente a Tableau Desktop, facilitando así el análisis visual posterior.



### 5.2.2. Implementación de visualizaciones comunes.

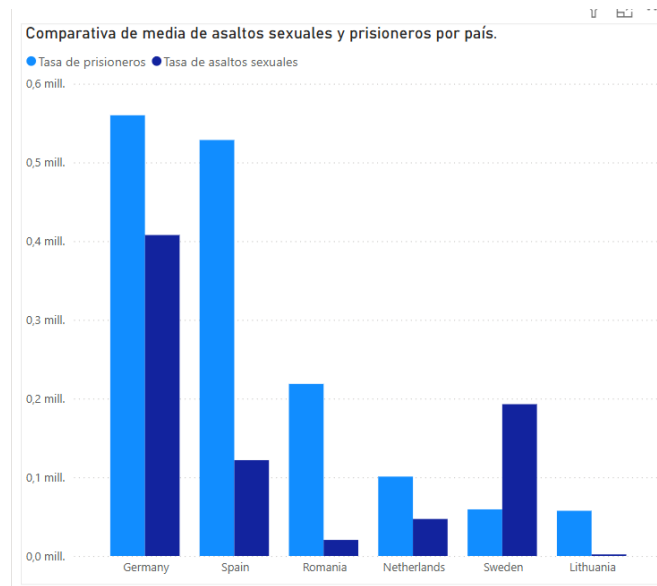
#### POWERBI

Una vez completada la limpieza de los datos y verificada la correcta relación entre las tablas, se ha procedido a la construcción de las visualizaciones en Power BI. Esta herramienta permite representar la información de forma dinámica e interactiva, facilitando la exploración de patrones y comparaciones entre países, años y tipos de delito.

La primera visualización desarrollada en Power BI ha sido un gráfico de barras agrupadas con el objetivo de comparar la tasa de asaltos sexuales y la tasa de población penitenciaria por país y año. En un primer intento se trató de incluir también los homicidios intencionados, pero dado que sus valores absolutos eran considerablemente más bajos, resultaban prácticamente imperceptibles en el gráfico, dificultando su interpretación visual.

Por esta razón, se optó por dividir la representación en dos gráficas de barras agrupadas, ubicadas dentro de un mismo dashboard para facilitar la comparación visual. La primera muestra la relación entre asaltos sexuales y número de prisioneros, mientras que la segunda se centra únicamente en la evolución de los homicidios. En la figura 23 y la figura 24 se puede observar el resultado de ambas visualizaciones, cada una destacando los valores con mayor claridad y permitiendo al usuario detectar posibles correlaciones o divergencias entre criminalidad y respuesta penal a lo largo del tiempo y entre los diferentes países analizados.

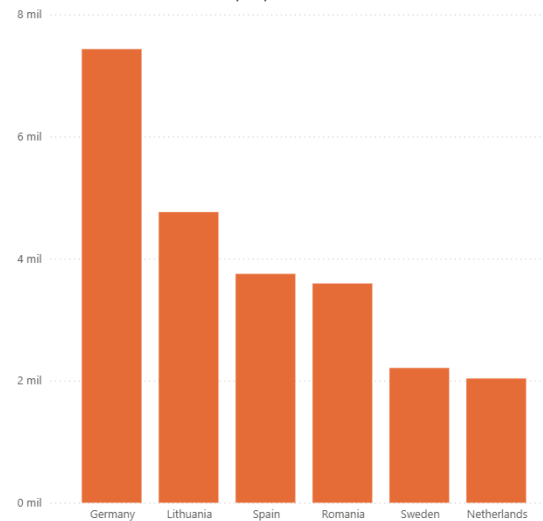
**Figura 23.**Gráfica 1 de barras agrupadas en PowerBI



Fuente: Elaboración propia.

**Figura 24.**Gráfica 2 de barras agrupadas en PowerBI

Media de la tasa de homicidios por país.

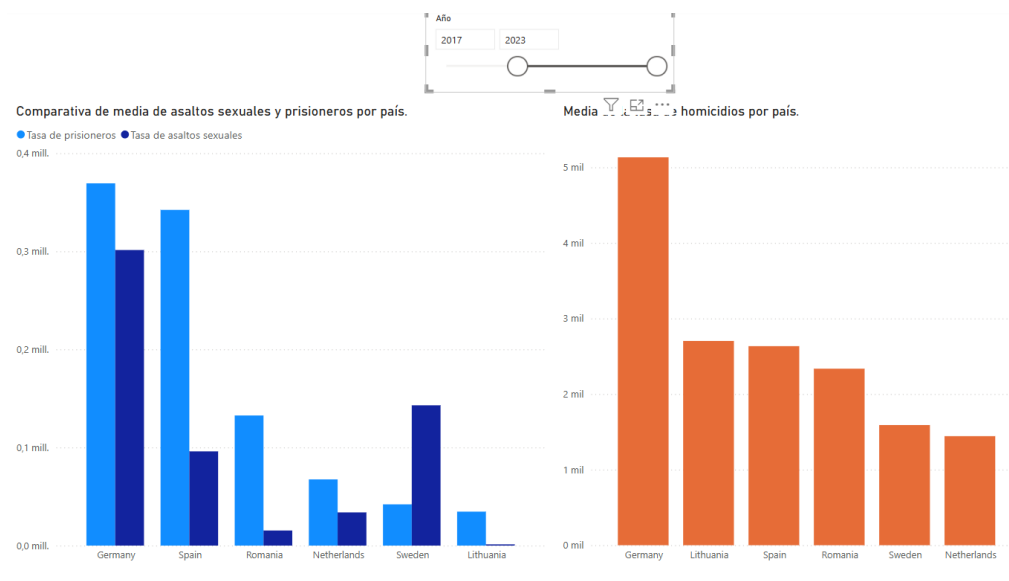


Fuente: Elaboración propia.

Además de ello, se ha incorporado un filtro por año, como se muestra en la figura 25, que permite al usuario seleccionar uno o varios años concretos para centrar la visualización en un periodo determinado. Esta funcionalidad, muy fácil de implementar en Power BI, facilita la comparación entre años específicos y permite observar la evolución temporal de cada variable de forma más detallada. No obstante, también es posible dejar activados todos los años para obtener una visión general de la década analizada, lo que resulta útil para detectar tendencias a largo plazo.

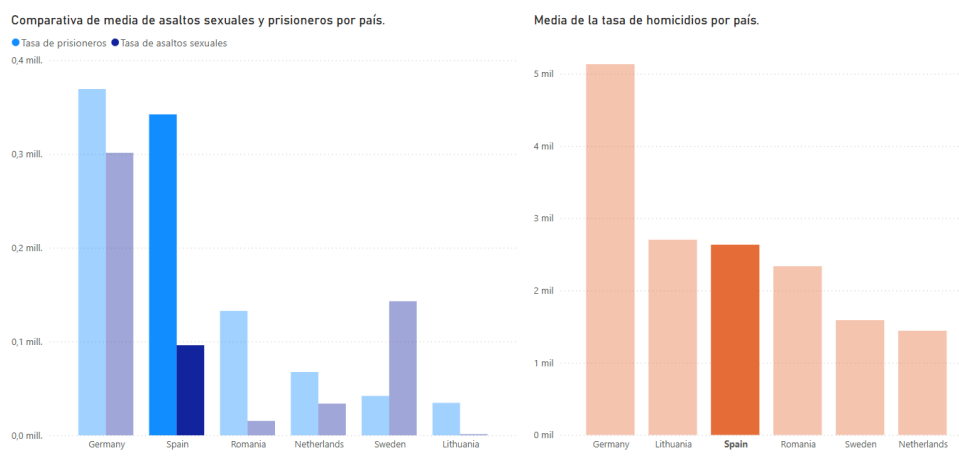
Power BI incorpora por defecto una funcionalidad muy útil para facilitar la exploración de los datos por parte del usuario. Cuando se pulsa sobre una barra en cualquiera de las dos gráficas del dashboard, se activa un efecto de resaltado (highlight), como podemos ver en la figura 26, que permite visualizar de forma inmediata la barra correspondiente al mismo país en la otra gráfica. Esta interacción cruzada mejora notablemente la experiencia del usuario, ya que permite detectar más fácilmente correspondencias o contrastes entre los distintos indicadores sin necesidad de aplicar filtros adicionales.

**Figura 25.** Highlight por defecto de PowerBI.



Fuente: Elaboración propia.

**Figura 26.** Filtro de año en el dashboard de barras agrupadas.

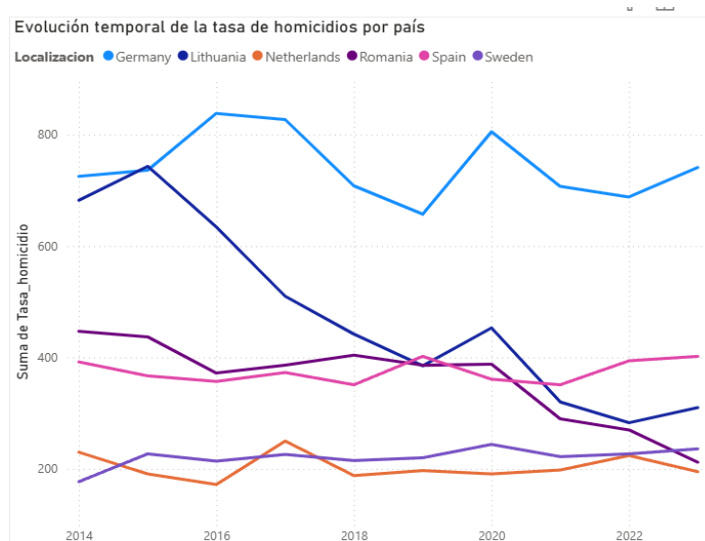


Fuente: Elaboración propia.

Otra de las visualizaciones desarrolladas en Power BI ha sido un gráfico de líneas que representa la evolución temporal de la tasa de homicidios por país. Esta gráfica permite observar con claridad la tendencia de este tipo de delito entre los años 2014 y 2022 en los seis países seleccionados.

Como se aprecia en la figura 27, se han trazado líneas separadas por país, lo que permite detectar de forma visual si las tasas han aumentado, disminuido o se han mantenido estables a lo largo del tiempo.

**Figura 27.** Gráfico de líneas en PowerBI.



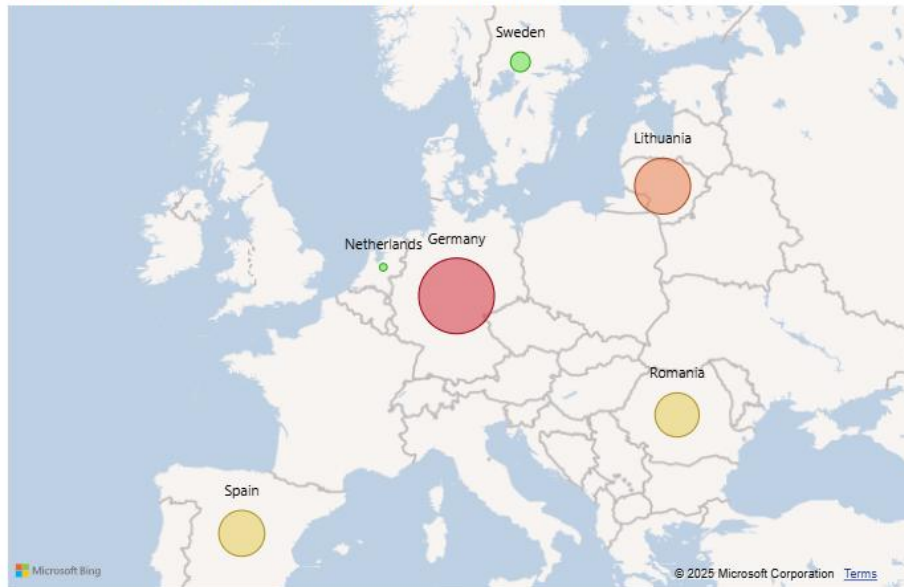
Fuente: Elaboración propia.

Este tipo de visualización resulta especialmente útil para identificar patrones en la criminalidad violenta y realizar comparaciones directas entre países y años. Además, Power BI permite al usuario interactuar con el gráfico al pasar el cursor por los puntos de cada línea, mostrando el valor exacto en cada año, lo que facilita un análisis más detallado sin necesidad de aplicar filtros adicionales.

Como complemento a las visualizaciones anteriores, se ha incluido un mapa de burbujas para representar de forma geográfica la tasa de homicidios por país. Esta visualización, que puede verse en la figura 28, permite identificar rápidamente las diferencias entre países en un solo vistazo, utilizando el tamaño y color de las burbujas para reflejar la magnitud del valor.

**Figura 28.** Mapa de burbujas en PowerBI.

Comparativa de la tasa de homicidios por localización



Fuente: Elaboración propia.

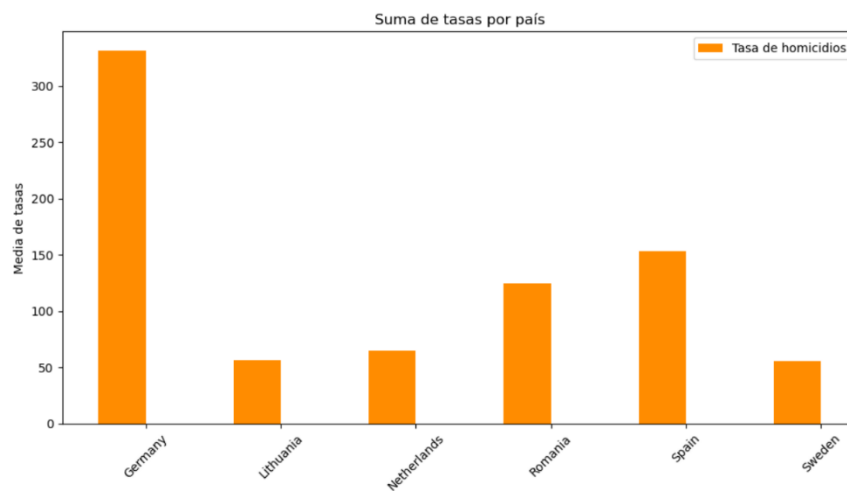
Cada burbuja está posicionada sobre el territorio correspondiente, y su tamaño representa el número total de homicidios, mientras que el color aporta un refuerzo visual para diferenciar los niveles más altos o bajos. Esta representación es especialmente útil para usuarios que buscan una interpretación espacial de los datos, pues facilita la localización y comparación directa de los países en el contexto europeo. Además, Power BI integra esta funcionalidad con Bing Maps, lo que permite generar automáticamente el posicionamiento geográfico a partir de los nombres de país incluidos en el dataset.

### PYTHON CON LIBRERÍAS.

Inicialmente se había previsto realizar las visualizaciones en Python utilizando únicamente la biblioteca Matplotlib, dada su popularidad y simplicidad para representar datos de forma estática. Sin embargo, tras generar los primeros gráficos, se observó que el nivel visual y funcional ofrecido por esta herramienta era demasiado básico en comparación con el de las plataformas de business intelligence utilizadas (Power BI y Tableau). Las gráficas creadas con Matplotlib, aunque correctas desde el punto de vista técnico, carecían de la interactividad y del atractivo visual necesarios para facilitar un análisis exploratorio fluido y comparable.

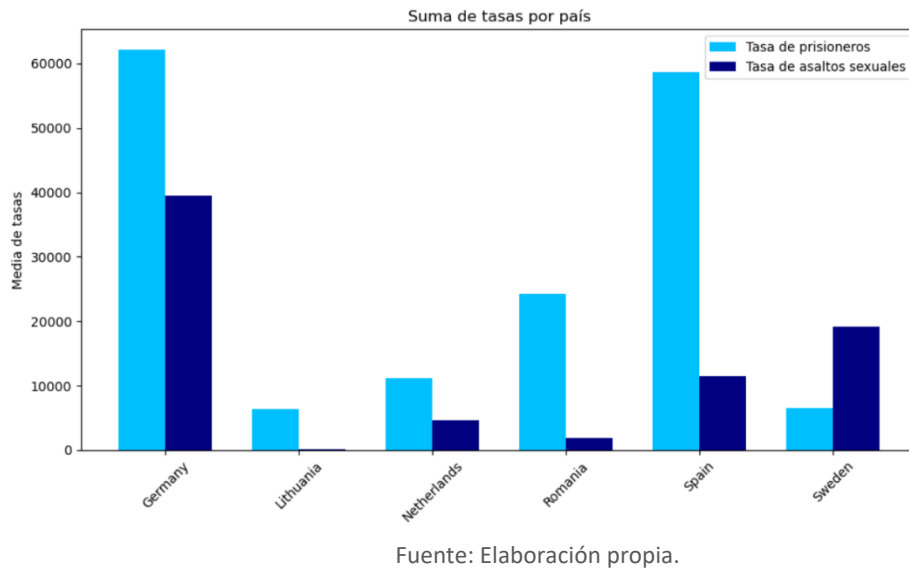
Por este motivo, se decidió complementar el trabajo en Python con la biblioteca Plotly, que permite generar visualizaciones más dinámicas e interactivas directamente desde el entorno de programación. Con Plotly, ha sido posible replicar las mismas gráficas construidas en Power BI, añadiendo funcionalidades como el filtrado por año, la visualización por categorías al pasar el ratón o la selección de elementos concretos. Esto permite al usuario interactuar con los datos de forma más flexible y mejora significativamente la experiencia visual sin renunciar al control y personalización que ofrece el código en Python.

**Figura 29.**Gráfico de columnas agrupadas con Python.



Fuente: Elaboración propia.

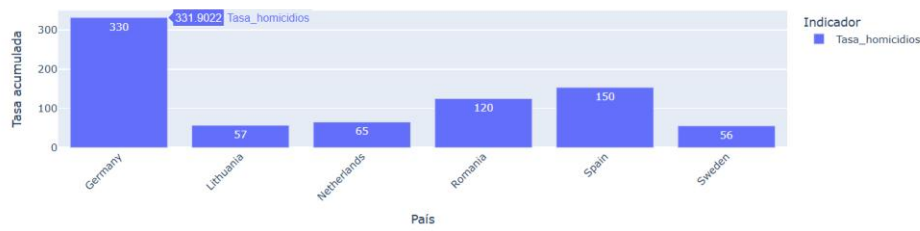
**Figura 30.** Grafico de columnas 2 en Python.



Las figuras 29 y 30 se muestran dos gráficos de columnas agrupadas elaborado con la biblioteca Matplotlib. Estas visualizaciones, tal y como se ha realizado en PowerBI, representan la media de las tasas de homicidios, asaltos sexuales y prisioneros entre los años 2014 y 2022 para cada uno de los seis países analizados. Aunque Matplotlib permite generar gráficos correctos desde el punto de vista técnico, su diseño por defecto es más limitado visualmente, con opciones reducidas de personalización y escasa interactividad para el usuario final. No se pueden aplicar filtros dinámicos ni explorar valores de forma detallada sin modificar directamente el código.

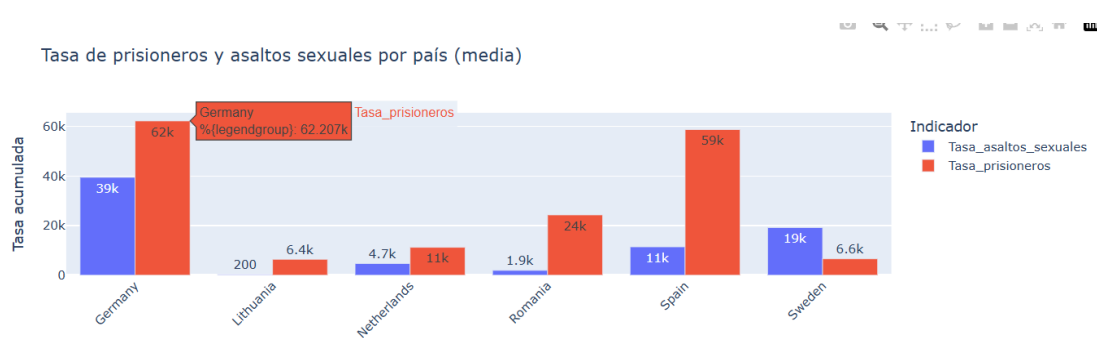
A pesar de estas limitaciones, esta visualización sirve como base para contrastar el tipo de resultados que pueden obtenerse en un entorno de programación tradicional, frente a las herramientas de business intelligence. Para mejorar la experiencia interactiva, esta misma gráfica ha sido replicada posteriormente utilizando Plotly.

Figura 31. Gráfica 1 Python con Plotly.



Fuente: Elaboración propia.

Figura 32. Gráfica 2 Python con Plotly.



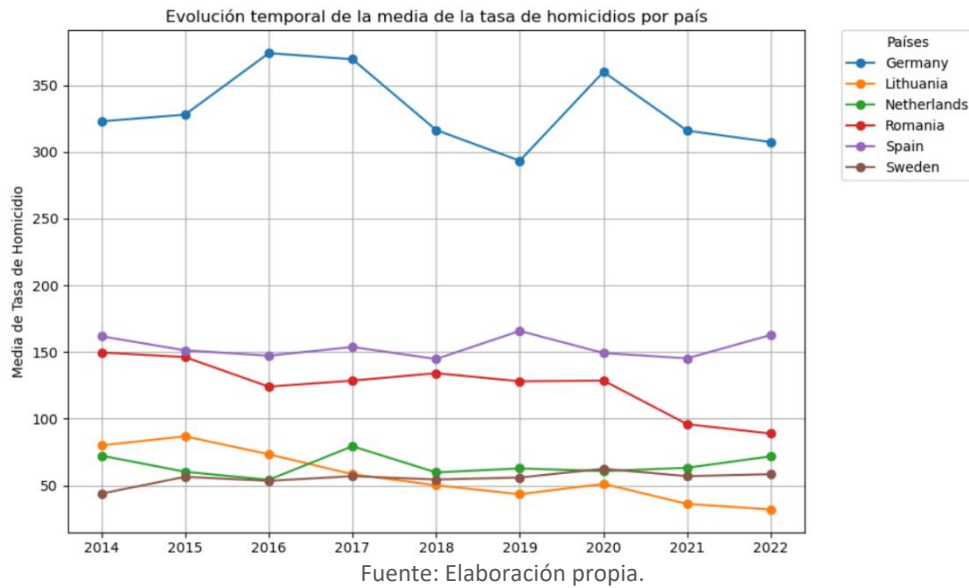
Fuente: Elaboración propia.

Las dos visualizaciones mostradas en las figuras 31 y 32 han sido generadas con la biblioteca Plotly y reproducen los mismos gráficos de columnas agrupadas desarrollados previamente en Matplotlib, pero con un diseño mucho más interactivo y visualmente atractivo. Gracias a las funcionalidades nativas de Plotly, estas gráficas permiten al usuario interactuar directamente con los elementos visuales, mostrando valores detallados al pasar el cursor por cada barra, y mejorando así la comprensión inmediata de los datos. Además, su integración con entornos como Jupyter Notebook facilita su uso dentro de flujos de análisis más amplios. Estas mejoras hacen que Plotly sea una alternativa muy válida frente a herramientas de BI, especialmente en contextos donde se requiera flexibilidad y personalización a través de código.

A continuación, y para seguir con la línea de PowerBI, se ha procedido a realizar la siguiente gráfica.



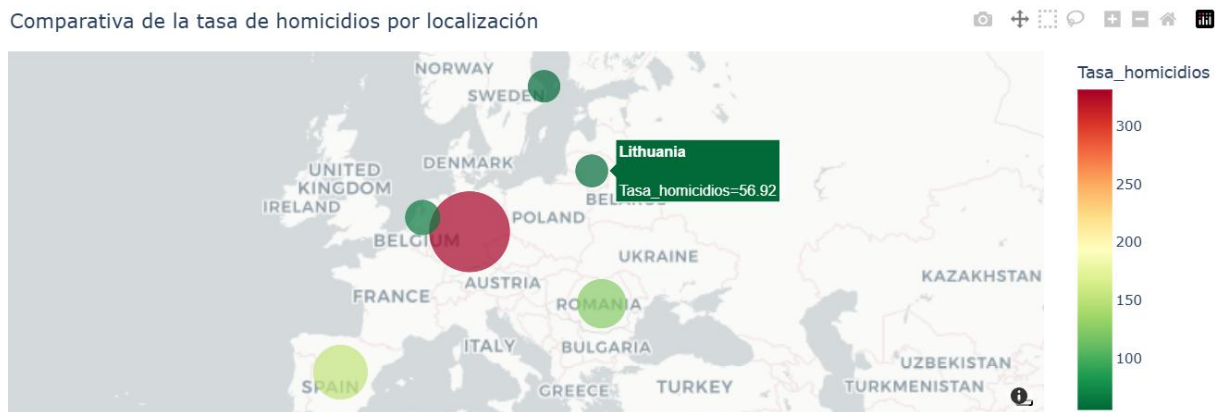
**Figura 33.**Gráfica evolución con Python.



En la figura 33 se muestra un gráfico de líneas desarrollado con Matplotlib. Cada línea corresponde a uno de los seis países analizados, permitiendo observar la tendencia que ha seguido este tipo de delito en cada caso. A pesar de la simplicidad visual característica de Matplotlib, este tipo de gráfico resulta útil para identificar cambios significativos a lo largo del tiempo y contrastar la estabilidad o variabilidad de las tasas entre países. En este caso, se ha optado solo por Matplotlib ya que su sintaxis es adecuada para representar series temporales, aunque sin interactividad.

**Figura 34.**Mapa de burbujas con Python.

Comparativa de la tasa de homicidios por localización



Fuente: Elaboración propia.

En la figura 34 se muestra un mapa de burbujas creado con Plotly que representa la tasa de homicidios por país en el contexto europeo. Este tipo de visualización no es posible con Matplotlib, ya que esta biblioteca no permite generar mapas de forma nativa. Por este motivo, se ha optado por utilizar Plotly, que sí ofrece soporte integrado para visualizaciones geográficas interactivas. No obstante, antes de generar el gráfico fue necesario añadir manualmente las coordenadas geográficas (latitud y longitud) de cada país, ya que Plotly no detecta automáticamente las ubicaciones a partir del nombre del país.

**Figura 35.**Dataset con coordenadas añadidas.

	País	Año	Tasa_homicidios	Tasa_asaltos_sexuales	Tasa_prisioneros	Latitud	Longitud
0	Germany	2014	645.0	34959	63228.0	51.1657	10.4515
1	Germany	2015	655.0	34265	63020.0	51.1657	10.4515
2	Germany	2016	747.0	37166	64291.0	51.1657	10.4515
3	Germany	2017	738.0	34815	65841.0	51.1657	10.4515
4	Germany	2018	632.0	40585	65762.0	51.1657	10.4515

Fuente: Elaboración propia.

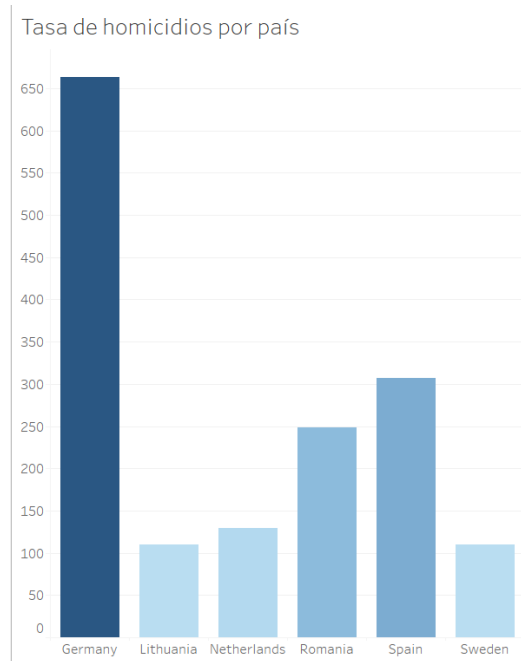
Para ello, se creó un diccionario con las coordenadas de cada una de las seis localizaciones analizadas y se incorporaron como nuevas columnas al dataframe, como se puede ver en la figura 35. El resultado permite visualizar cada país con una burbuja cuyo tamaño y color reflejan la tasa de homicidios, incluyendo además una escala de colores que facilita la interpretación de los valores. Al pasar el cursor por cada burbuja, se muestran los valores detallados de cada país, lo que mejora la experiencia de análisis en comparación con los gráficos estáticos.

### TABLEAU.

Tras la preparación y limpieza de los datos en Tableau Prep, se han generado diversas visualizaciones en Tableau Desktop con el objetivo de representar de forma clara y comparativa los principales indicadores analizados. Aunque Tableau presenta algunas limitaciones en cuanto a interactividad avanzada respecto a otras herramientas, destaca por

su claridad visual, su capacidad de generar dashboards limpios y su rápida integración con los datos preparados previamente.

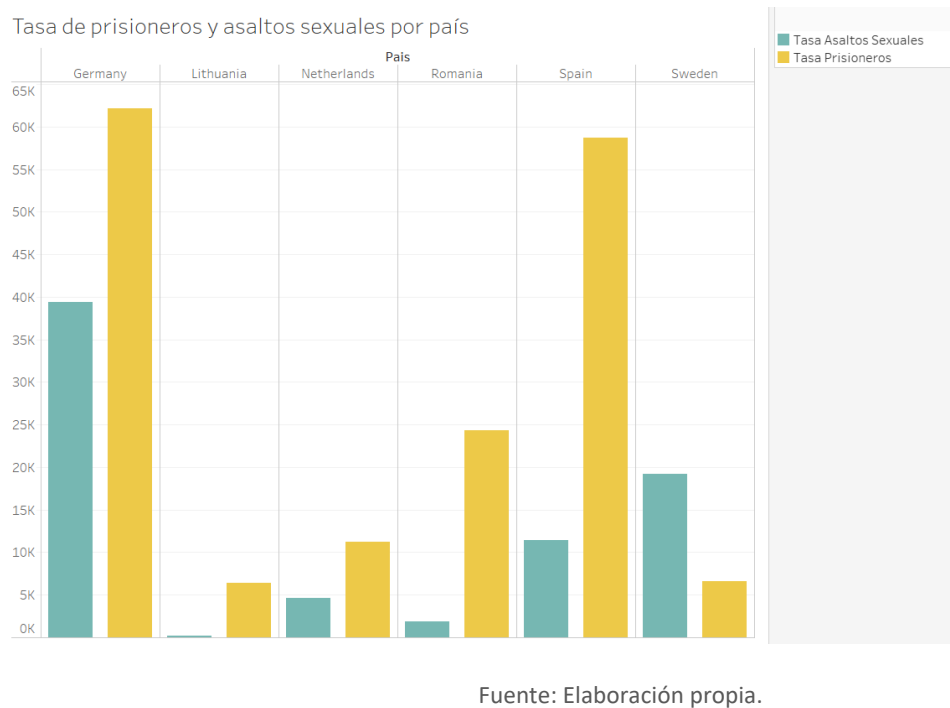
**Figura 36. Barras agrupadas en Tableau.**



Fuente: Elaboración propia.

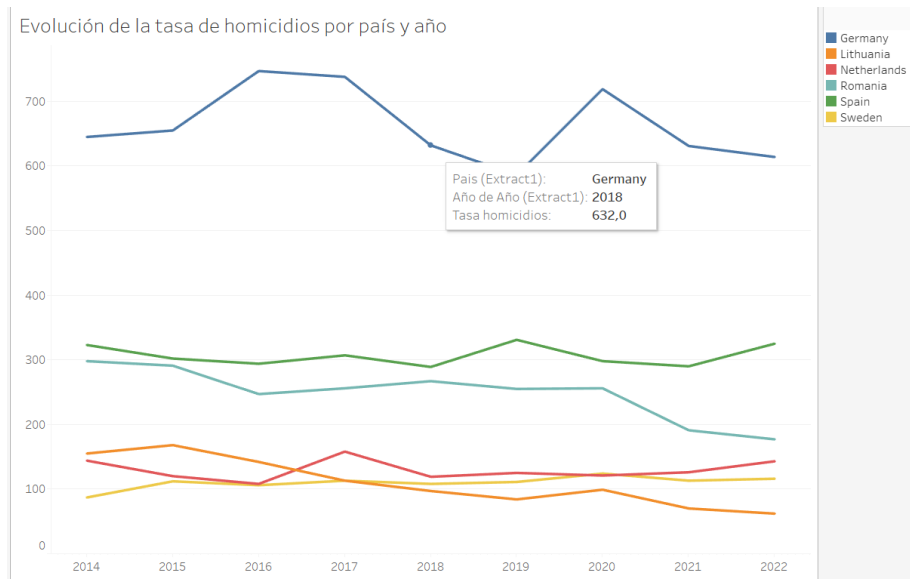
En la figura 36 se presenta el gráfico de barras agrupadas que muestra la tasa de homicidios por país, utilizando un esquema de color progresivo para reforzar visualmente las diferencias entre los valores. En esta herramienta no se genera el efecto resaltado automático como ocurría en PowerBI.

**Figura 37.** Barras agrupadas 2 en Tableau



Por otro lado, la figura 37 muestra otro gráfico de barras agrupadas en el que se representan conjuntamente la tasa de prisioneros y la tasa de asaltos sexuales por país. La elección de una paleta diferenciada para cada indicador facilita la comparación directa y la cuadrícula vertical por defecto segmenta visualmente cada país, ayudando a interpretar la información. La leyenda aparece desplazada fuera del área principal del gráfico y no puede integrarse directamente dentro, queda situada en un panel lateral con fondo gris, lo cual desentona ligeramente con el diseño general y rompe la armonía visual de la composición.

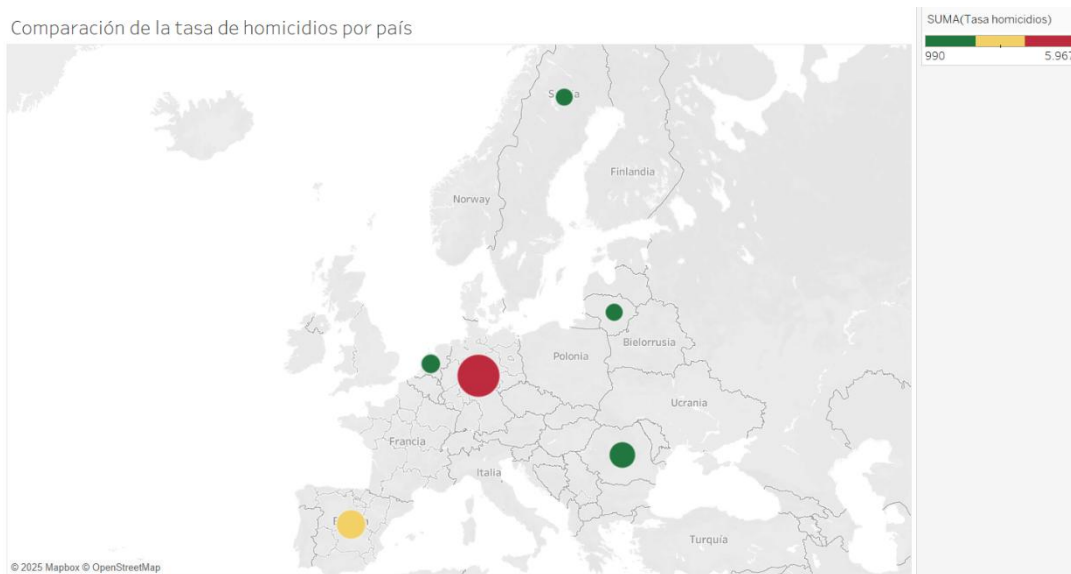
**Figura 38.**Gráfica evolución Tableau.



Fuente: Elaboración propia.

En la figura 38 se presenta el gráfico de líneas generado para representar la evolución de la tasa de homicidios por país. A diferencia de las versiones creadas en Python con Matplotlib o Plotly, esta visualización destaca por su limpieza visual y claridad en la presentación. La paleta de colores es sencilla y efectiva. Sin embargo, la leyenda, que nuevamente queda situada en el lateral, no puede integrarse dentro del gráfico, lo que rompe ligeramente la armonía visual. Aun así, la fluidez al interactuar con los datos y la capacidad de mostrar información al pasar el cursor por cada punto refuerzan su utilidad, manteniendo una buena experiencia de análisis sin necesidad de código, al igual que ocurre con PowerBI.

**Figura 39.** Mapa de burbujas en Tableau.



Fuente: Elaboración propia.

En la figura 39 se muestra el mapa de burbujas generado para representar la tasa de homicidios por país. Visualmente, el resultado es limpio y claro, con una buena integración del mapa base y una leyenda cromática que refuerza la lectura de los valores mediante un gradiente de color de verde a rojo. El mapa base utilizado en la visualización corresponde a una integración automática con OpenStreetMap, que permite representar los datos geográficos sin necesidad de configuraciones adicionales. En comparación con los mapas realizados en Plotly, Tableau ofrece una solución equilibrada: no requiere configurar manualmente las coordenadas geográficas y posiciona las burbujas correctamente sobre el mapa. De nuevo, la leyenda no se ha podido integrar en el área principal del gráfico.

Un aspecto interesante de Tableau, es que ha creado medidas y variables automáticas (mostradas en la figura 40) por defecto, algunas de las cuales han sido de gran ayuda, como las de coordenadas o la de valores de medidas, utilizada en la primera gráfica.

**Figura 40.** Variables creadas automáticamente en Tableau.

Abc *Nombres de medidas*

---

🌐 *Latitud (generado)*

🌐 *Longitud (generado)*

# *Valores de medidas*

Fuente: Elaboración propia.

### 5.2.3. Evaluación técnica de cada herramienta

Desde el punto de vista técnico, Power BI ha mostrado un rendimiento general bastante fluido, con tiempos de carga rápidos tanto al importar los datasets como al navegar por los dashboards. La respuesta a los filtros aplicados, como las segmentaciones por año o país, ha sido inmediata incluso al trabajar con múltiples visualizaciones conectadas dentro de una misma página. Entre las limitaciones técnicas encontradas destaca la necesidad de estructurar bien las relaciones entre tablas antes de iniciar el análisis, ya que una mala configuración del modelo de datos puede provocar errores en los resultados mostrados. En cuanto a la documentación y soporte, Power BI cuenta con una amplia base de recursos oficiales, foros comunitarios activos y tutoriales, lo que facilita la resolución de dudas, aunque en algunos casos concretos la búsqueda de soluciones específicas puede requerir cierto tiempo por la dispersión de contenidos.

En el caso de Python, el rendimiento ha sido muy sólido, especialmente al trabajar con la biblioteca Pandas para la carga y manipulación de los datos. Los tiempos de procesamiento han sido rápidos incluso con la integración de múltiples datasets, y las operaciones de filtrado o agrupación se ejecutan de forma eficiente siempre que el código esté optimizado. Sin embargo, al tratarse de un entorno de programación, la respuesta a filtros o visualizaciones no es interactiva por defecto, lo que limita la experiencia del usuario final si no se incorporan librerías adicionales como Plotly. Entre las limitaciones técnicas destaca la necesidad de preparar manualmente ciertas tareas que en herramientas de BI son automáticas, como la obtención de coordenadas geográficas o la estructuración visual del modelo de datos. En cuanto a documentación y soporte, Python cuenta con una comunidad muy activa y una extensa cantidad de recursos, aunque la calidad y claridad de las soluciones puede variar según la fuente y el nivel técnico del usuario.

En términos de rendimiento, Tableau ha mostrado un comportamiento dispar entre sus dos entornos principales. Por un lado, Tableau Prep ha ofrecido un tiempo de carga bastante rápido durante el proceso de limpieza y transformación de los datos, permitiendo aplicar filtros, renombrar columnas y realizar uniones de forma ágil. Sin embargo, al pasar los datos a Tableau Desktop, los tiempos de carga y respuesta han sido notablemente más lentos, especialmente al trabajar con múltiples visualizaciones en un mismo dashboard o al aplicar filtros cruzados. Esta diferencia puede dificultar el trabajo cuando se requiere fluidez en la

exploración visual. En cuanto a las limitaciones técnicas, Tableau no permite crear modelos de datos complejos con relaciones múltiples como en Power BI, lo que obliga a preparar los datos previamente en estructuras planas o realizar uniones manuales. Por otro lado, la documentación y el soporte oficial son sólidos, con una comunidad activa y recursos bien organizados, aunque algunas funciones avanzadas requieren cierta curva de aprendizaje para usuarios sin experiencia previa.

### 5.3.DISCUSIÓN DE ANÁLISIS Y RESULTADOS.

#### 5.3.1. Ventajas y desventajas observadas

Para sintetizar los resultados obtenidos en la evaluación de las tres herramientas analizadas, se ha elaborado una tabla comparativa basada en los criterios definidos previamente en la fase de evaluación: facilidad de uso, calidad visual, rendimiento, flexibilidad de análisis y curva de aprendizaje. Esta tabla (Tabla 1) permite observar de forma clara y estructurada las ventajas y desventajas detectadas en cada caso, facilitando así una valoración global de las soluciones en función de sus características técnicas y funcionales. La comparación se ha realizado a partir de la experiencia directa en el uso de cada herramienta y tomando como referencia tanto su comportamiento durante el desarrollo como las posibilidades que ofrecen para proyectos similares.

**Tabla 1.Comparación de criterios entre herramientas.**

Criterio	PowerBI	Python(Matplotlib/Plotly)	Tableau
Facilidad de uso	Alta: interfaz muy intuitiva y con muchas funcionalidades guiadas.	Baja: requiere conocimientos de programación previos.	Media: interfaz clara pero menos intuitiva que PowerBI
Calidad Visual	Alta: Visualizaciones limpias y atractivas por defecto. Muchos filtros automáticos.	Variable: limitada en Matplotlib y media en Plotly. Aunque en Plotly se pueden añadir interacciones, no es comparable con herramientas BI.	Media: diseño limpio y profesional por defecto. Mayor dificultad para añadir filtros.



Rendimiento	Muy bueno: carga rápida y respuesta fluida a los filtros	Muy buena: procesamiento rápido con un código optimizado	Dispar: un poco más rápido en Prep, más lento en Desktop.
Flexibilidad de análisis	Media: admite personalización limitada sin DAX avanzado.	Alta: permite máxima personalización mediante código.	Media: buena para análisis directo, limitada en el modelo de datos.
Curva de aprendizaje	Baja, se aprende rápido, ideal para perfiles no técnicos.	Alta: requiere experiencia previa en programación.	Media: curva moderada, aunque la interfaz parece fácil, requiere práctica para tener más fluidez.

Fuente: Elaboración propia.

### 5.3.2. Casos de uso más adecuados para cada solución.

Power BI es especialmente recomendable para analistas sin conocimientos de programación o para profesionales que necesitan obtener resultados visuales rápidamente sin dedicar demasiado tiempo al aprendizaje de la herramienta. Su integración con otras aplicaciones de Microsoft, su interfaz intuitiva y su buena respuesta a la interacción lo convierten en una opción ideal para entornos empresariales o administrativos, incluidos los del sector público, donde se requieren informes claros, automatizables y fáciles de mantener. Además, es una buena alternativa para equipos multidisciplinarios en los que se necesita compartir dashboards de forma sencilla.

Python, con bibliotecas como Pandas, Matplotlib y Plotly, es la herramienta más adecuada para investigadores académicos, científicos de datos o perfiles técnicos avanzados que buscan el máximo control sobre el análisis y la visualización. Su principal ventaja reside en la flexibilidad, que permite adaptar el análisis a cualquier necesidad específica, automatizar procesos y trabajar con volúmenes de datos más complejos. Aunque requiere conocimientos de programación, es una solución muy potente para generar visualizaciones personalizadas e integrarlas en entornos de análisis más amplios.

Tableau se sitúa en un punto intermedio entre las dos anteriores. Es una herramienta adecuada para profesionales del análisis de datos que valoran la estética y la claridad visual, sin necesidad de entrar en programación ni depender de ecosistemas concretos como el de Microsoft. También puede ser útil en entornos institucionales o de consultoría, donde se requiere generar dashboards atractivos para presentar resultados a públicos no especializados. Si bien su curva de aprendizaje inicial puede ser algo más lenta que la de Power BI, ofrece una buena experiencia una vez dominada.

### 5.3.3. Consideraciones finales sobre la aplicabilidad.

El uso de herramientas de visualización como Power BI, Tableau o Python con librerías gráficas representa un recurso clave para facilitar la toma de decisiones basada en datos, especialmente en ámbitos complejos como el análisis de la criminalidad. La capacidad de transformar grandes volúmenes de información en gráficos claros e interpretables permite identificar patrones, evaluar políticas públicas y orientar estrategias. De cara al futuro, una posible línea de mejora sería la integración de técnicas de análisis predictivo o aprendizaje automático en estas plataformas, así como la exploración de herramientas más especializadas en visual analytics para contextos criminológicos o judiciales. También podría estudiarse la usabilidad de estas herramientas por parte de usuarios no expertos, lo que permitiría ampliar su impacto en contextos institucionales.

## 6. Código fuente y datos analizados

### 6.1. Código fuente

Todo el código fuente desarrollado en el marco de este Trabajo Fin de Máster está alojado en un repositorio de GitHub gestionado exclusivamente por el autor de este TFM. El repositorio puede consultarse en la siguiente URL:

<https://github.com/susibrg/TFMRepository>

En él se incluye la totalidad de los scripts, notebooks y ficheros de configuración utilizados para la extracción, limpieza, análisis y visualización de los datos. El historial de commits refleja únicamente contribuciones realizadas por el autor, sin participación de ningún otro usuario, lo que garantiza la autoría y titularidad íntegra del código.

## 6.2. Datos Analizados

Asimismo, los datos empleados para el análisis proceden de Eurostat y se encuentran incorporados directamente en el repositorio, organizados en los propios scripts de limpieza y en los notebooks o ficheros que generan cada uno de los gráficos para Power BI, Tableau y Python. De este modo, cualquier persona que acceda al repositorio podrá reproducir íntegramente tanto la obtención de los datos como el procesamiento y la generación de las visualizaciones.

## 7. Conclusiones

En este Trabajo Fin de Máster se ha abordado el reto de comparar tres herramientas de visualización de datos—Power BI, Tableau y Python (Matplotlib/Plotly)—aplicadas al análisis de datos de criminalidad en Europa, utilizando como fuente primaria los conjuntos de datos agregados y anonimizados de Eurostat. Partiendo de la necesidad de elegir la solución más adecuada según criterios de rendimiento, facilidad de uso, flexibilidad y adecuación a distintos perfiles de usuario, se definió una metodología basada en:

- Preparación y limpieza de los datos.
- Implementación de visualizaciones de referencia en cada herramienta.
- Evaluación técnica y funcional (rendimiento, tiempos de carga, respuesta a filtros, limitaciones).
- Comparativa de ventajas y desventajas.
- Recomendaciones de uso según perfil y casos prácticos.

Los principales logros y aportaciones de este trabajo son:

- Validación del rendimiento: se confirmó que Python ofrece la máxima flexibilidad, aunque requiere mayor esfuerzo de optimización; Power BI destaca por su velocidad de carga en entornos empresariales; y Tableau demuestra un rendimiento sólido en preparación de datos (Prep) aunque algo más lento en escritorio.
- Análisis de usabilidad: Power BI resulta intuitivo para analistas sin código, Tableau favorece la exploración visual sin programación, y Python es idóneo para perfiles técnicos capaces de controlar cada detalle.

- Reproducibilidad y transparencia: todo el código y los datos de Eurostat están disponibles en el repositorio del autor, garantizando trazabilidad de los procesos y resultados.

**Tabla 2.** Comparación de objetivos planteados con resultados obtenidos.

Objetivo	Conclusión principal
Identificar las características clave de cada herramienta de visualización (Tableau, Power BI y Python con Plotly/Matplotlib) en relación con su aplicabilidad al análisis de datos criminalísticos.	Se han detallado las funciones, extensiones y posibilidades de interfaz de usuario de cada herramienta en el contexto de análisis criminalístico, destacando sus fortalezas técnicas y limitaciones operativas.
Comparar la facilidad de uso, el rendimiento y la calidad de las visualizaciones ofrecidas por cada herramienta.	La evaluación mostró que Power BI y Tableau ofrecen interfaces más intuitivas y tiempos de carga reducidos, mientras que Python requiere mayor configuración, pero permite un control fino de la estética y la interactividad de los gráficos.
Desarrollar visualizaciones representativas a partir de un conjunto de datos criminalísticos europeos.	Se generaron ejemplos comparativos — columnas, líneas, burbujas y mapas— en las tres plataformas, utilizando los mismos scripts y parámetros de diseño para garantizar la homogeneidad en la interpretación de los resultados.
Evaluar la efectividad de cada herramienta en la interpretación y comunicación de tendencias delictivas	Se verificó que Tableau facilita la exploración visual “out-of-the-box”, Power BI integra filtros rápidos y Python destaca en la personalización avanzada de anotaciones y dinámicas.

Valorar si cada herramienta puede ser útil para instituciones públicas o entidades que analizan la criminalidad.	Se definieron escenarios de uso para departamentos gubernamentales y agencias de seguridad, concluyendo que Power BI y Tableau son más adecuados para informes periódicos y Python para proyectos de investigación académica o desarrollo ad hoc.
Resumir los resultados y proponer qué herramienta es más recomendable según el tipo de usuario o situación	Se elaboró una matriz de recomendaciones finales, ofreciendo prácticas para su implementación en proyectos reales de análisis de criminalidad.

Fuente: Elaboración propia

En conjunto, este estudio confirma que no existe una única “mejor” solución, sino que la selección debe adaptarse al contexto y al perfil del usuario. Las aportaciones de este TFM proporcionan una guía fundamentada que facilita la toma de decisiones en proyectos de análisis de criminalidad, así como una base para futuras investigaciones en automatización de pipelines y optimización de entornos de Big Data.

## 8. Limitaciones y prospectiva

### 8.1. Limitaciones

En este TFM se siguió un proceso claro: limpieza y preparación de los datos, creación de gráficos comparativos en Power BI, Tableau y Python, y evaluación del funcionamiento de cada herramienta. Este enfoque permitió identificar ventajas y limitaciones, si bien no se incorporaron criterios adicionales como pruebas de usabilidad con usuarios o encuestas de opinión. También se emplearon versiones específicas de librerías de Python, lo que podría provocar diferencias en la reproducción del trabajo si se utilizan versiones distintas.

Los datos provienen únicamente de Eurostat en formato de tablas agregadas, lo que impidió acceder a información desagregada por edad, género u otros factores. Tampoco se consideraron delitos como la ciberdelincuencia o la violencia doméstica, ni se introdujeron indicadores económicos o sociales que habrían aportado mayor contexto a las cifras de

criminalidad. Al centrarse en homicidios, agresiones sexuales y robos hasta 2022, quedó fuera el análisis de tendencias más recientes y de relaciones más complejas entre variables.

El análisis de rendimiento se realizó en una única máquina con características determinadas, por lo que los tiempos de carga y respuesta pueden variar en otros entornos o servidores más potentes. Tampoco se midió el consumo de recursos en escenarios de datos de gran volumen, ni se exploraron visualizaciones avanzadas como diagramas de flujo o paneles interactivos complejos. Estas limitaciones abren la puerta a futuros trabajos que amplíen la variedad de datos, experimenten con nuevas visualizaciones o incluyan estudios de usabilidad con usuarios reales.

## 8.2. Trabajo futuro

El análisis comparativo realizado abre diversas oportunidades para profundizar en el estudio de la criminalidad mediante visualizaciones más completas y adaptadas a contextos específicos. En primer lugar, sería beneficioso incorporar microdatos anonimizados que permitan un análisis más variado y segmentado, por ejemplo, por edad, género o nivel socioeconómico. De este modo, se podría explorar con mayor detalle cómo varían las tasas de delincuencia dentro de subgrupos poblacionales y enriquecer las conclusiones derivadas de los datos agregados de Eurostat.

Además, ampliar el conjunto de variables con indicadores socioeconómicos y de gobernanza —como el gasto público en seguridad o las tasas de pobreza— aportaría un contexto más amplio para interpretar las tendencias delictivas. La integración de estas variables en modelos de correlación o regresión permitiría identificar posibles relaciones causales y mejorar la comprensión de los factores que influyen en la evolución de la criminalidad.

Otra línea de trabajo futuro consiste en la realización de estudios de usabilidad con usuarios representativos de los perfiles objetivo (analistas de negocio, investigadores académicos, responsables de políticas públicas). Mediante pruebas con grupos o encuestas estructuradas, se podría validar empíricamente la percepción de facilidad de uso, la eficiencia en la creación de visualizaciones y la calidad de la comunicación de resultados, ajustando así las recomendaciones a las necesidades reales de los usuarios.

En el plano técnico, evaluar el rendimiento en entornos de computación distribuida —tales como clusters de alto rendimiento o servicios en la nube— y medir la escalabilidad ante

volúmenes de datos muy superiores permitiría trasladar este trabajo al ámbito de Big Data. Este enfoque ayudaría a determinar cómo optimizar los pipelines de procesamiento y a comparar la eficiencia de cada herramienta cuando se enfrentan a grandes conjuntos de datos en entornos productivos.

Finalmente, explorar otras librerías y plataformas de visualización, como D3.js para desarrollos web avanzados, Power BI Embedded para soluciones integradas en aplicaciones empresariales o software GIS para análisis espacial detallado, ampliaría el alcance de la comparativa. Estos entornos especializados podrían aportar valor añadido al ofrecer nuevas formas de interacción, visualizaciones geoespaciales o dashboards embebidos que faciliten la toma de decisiones en organizaciones públicas y privadas dedicadas al estudio y la prevención de la criminalidad.

## Referencias bibliográficas

AceleraPyme. (s.f.). Kit Digital. <https://www.acelerapyme.gob.es/kit-digital>

Anzilotti, E. (2021, 24 de junio). Police Scorecard visualizes data on policing to drive equity and change. Tableau. <https://www.tableau.com/blog/police-scorecard-visualizes-data-policing-drive-equity-and-change>

Apache Superset. (s.f.). Apache Superset. <https://superset.apache.org/>

Cairo, A. (2012). The functional art: An introduction to information graphics and visualization. New Riders.

Cairo, A. (2021, 16 de agosto). My Favorite Tools: Alberto Cairo on Data Visualization. Global Investigative Journalism Network. <https://giin.org/stories/my-favorite-tools-alberto-cairo-on-data-visualization/>

Cartoy García, P. (2023). *Guía de uso de Eurostat*. Comisión Europea.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, 19(2), 171–209. <https://doi.org/10.1007/s11036-013-0489-0>

Esri. (s.f.). Representación cartográfica y análisis de la delincuencia. Recuperado de <https://www.esri.com/es-es/industries/law-enforcement/strategies/crime-analysis>

European Union Agency for Fundamental Rights (FRA). (s.f.). Data and maps. <https://fra.europa.eu/en/publications-and-resources/data-and-maps>

Eurostat. (2024). Crime statistics - Statistics Explained. Recuperado de [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Crime\\_statistics](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Crime_statistics)

Eurostat. (s.f.). Crime and criminal justice statistics. <https://ec.europa.eu/eurostat/web/crime/data/database>

Few, S. (2009). Now you see it: Simple visualization techniques for quantitative analysis. Analytics Press.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>



IBM. (n.d.). *CRISP-DM help overview*. IBM.

<https://www.ibm.com/docs/es/spss-modeler/saas?topic=dm-crisp-help-overview>

Igarapé Institute. (s.f.). CrimeRadar. <https://igarape.org.br/en/tech/crimeradar/>

Kadam, A. J., & Akhade, K. (2021). A review on comparative study of popular data visualization tools. *Alochana Chakra Journal*, 10(4), 2916–2922. <https://alochana.org/wp-content/uploads/53-AJ2368.pdf>

Kumar, A. (2021). How Power BI can be used in a police force. LinkedIn.

<https://www.linkedin.com/pulse/how-powerbi-can-used-police-force-anurodh-kumar-aqudf>

Lieberman, S. (2016). Trying Out the NYPD's Yelp for Crime. *New York Magazine*. Recuperado de <https://nymag.com/intelligencer/2016/02/trying-out-the-nypds-yelp-for-crime.html>

McKinney, W. (2012). *Python for Data Analysis*. O'Reilly Media.

Microsoft Power BI. (s.f.). Business intelligence like never before.

<https://powerbi.microsoft.com/>

Muggah, R. (2017, 2 de febrero). O que acontece quando podemos prever crimes antes que eles aconteçam? Instituto Igarapé. <https://igarape.org.br/o-que-acontece-quando-podemos-prever-crimes-antes-que-eles-acontecam/>

NYPD. (s.f.). CompStat 2.0. New York City Police Department.

<https://compstat.nypdonline.org>

Parthe, R. M. (2023). Comparative Analysis of Data Visualization Tools: Power BI and Tableau. ResearchGate.

[https://www.researchgate.net/publication/374957892\\_Comparative\\_Analysis\\_of\\_Data\\_Visualization\\_Tools\\_Power\\_BI\\_and\\_Tableau](https://www.researchgate.net/publication/374957892_Comparative_Analysis_of_Data_Visualization_Tools_Power_BI_and_Tableau)

Pavlenko, V. (2022). Power BI Pros and Cons. AltexSoft.

<https://www.altexsoft.com/blog/power-bi-pros-cons/>

Plan de Recuperación. (2024, 27 de febrero). El programa Kit Digital llega a 530.000 beneficiarios, el 80% del objetivo comprometido con Bruselas.

<https://planderecuperacion.gob.es/noticias/programa-kit-digital-llega-530000-beneficiarios-80-por-ciento-objetivo-comprometido-Bruselas-prtr>

Plotly. (s.f.). Plotly Python Graphing Library. <https://plotly.com/python/>

QMetrix. (s.f.). Strengths and Limitations of Power BI. <https://qmetrix.com.au/strengths-and-limitations-power-bi/>

Raj, A. (2022, November 2). *Data-Ink Ratio Explained With Example*. CodeConquest. <https://www.codeconquest.com/blog/data-ink-ratio-explained-with-example/>

Ribas, J. I., Capllonch, F., & Prado, J. (2020). Análisis inteligente de datos y visualización aplicadas a la investigación criminal. Universidad Nacional del Sur. Recuperado de <https://repositoriodigital.uns.edu.ar/handle/123456789/6848>

Tableau. (s.f.). Tableau: Visual analytics platform. <https://www.tableau.com/>

Thinklytics. (s.f.). What are the Pros & Cons of Tableau?. <https://thinklytics.com/what-are-the-pros-cons-of-tableau/>

Tufte, E. R. (2001). *The visual display of quantitative information* (2nd ed.). Graphics Press.

Unión Europea. (2016). Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos. *Diario Oficial de la Unión Europea*.

United Nations Office on Drugs and Crime (UNODC). (s.f.). UNODC Data Portal. <https://dataunodc.un.org/>

Urabayan, M. (2002). Algunos antecedentes históricos de la infografía moderna. Infografía. I Exposición de Gráficos Periodísticos. Pamplona: SND-E.

Wheeler, A. (2021, 15 de marzo). Crime analysis dashboards in Tableau. <https://andrewpwheeler.com/2021/03/15/crime-analysis-dashboards-in-tableau/>