

# El Arte de la Analítica Criminal

Dossier de exploración Analítica mediante Técnicas y Herramientas de Excel.

Rodrigo Daniel Pastén Turén.  
Sgto. 1° de Carabineros (R) de Orden y Seguridad  
Magister en Data Science  
Doctorando en Data Science

© 2024 Autor. Todos los derechos reservados.

Se autoriza la reproducción total o parcial, solo para fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

## Palabras del autor

Estimados Analistas Criminales,

Nos encontramos en un escenario urbano que está en constante evolución, un entorno dinámico donde el desafío del crimen se transforma día a día. En este contexto, la labor que ustedes desempeñan o están a punto de emprender, ya sea en Carabineros de Chile, Policía de Investigaciones de Chile o en las áreas de gestión de Seguridad Municipales, es más crucial que nunca. Son ustedes los que, con su aguda visión y análisis meticuloso, desempeñan un papel fundamental en la prevención y solución de delitos, contribuyendo significativamente a la seguridad de nuestra sociedad.

Su rol va más allá de la mera observación y análisis de datos. Ustedes son los estrategas en la sombra, los artífices de tácticas proactivas y preventivas, que nos permiten no solo entender, sino también anticipar y responder eficazmente a los patrones delictuales. La combinación única de sus habilidades analíticas y conocimientos en criminología es esencial para mantener a nuestras comunidades seguras y pacíficas.

Esta guía que tienen ante ustedes es una fuente invaluable de conocimiento y herramientas. Esta no solo enriquecerá su comprensión teórica, sino que también agudizará sus habilidades analíticas para una aplicación práctica en el desafiante y fascinante mundo del análisis criminal, permitiendo a los gestores en seguridad determinar planificaciones un análisis más más efectivas.

El objetivo de este documento es proporcionarles un camino claro hacia la maestría en su campo, alineando sus habilidades con las altas demandas y expectativas de esta noble profesión. Les insto a aprovechar al máximo esta oportunidad de crecimiento profesional, a compartir sus percepciones, a desafiar el statu quo y a contribuir activamente a nuestra misión colectiva de crear comunidades más seguras.

Su expertise y dedicación son el baluarte que nos permite enfrentar con confianza los retos delictuales urbanos contemporáneos. Juntos, con pasión y compromiso, continuaremos trabajando para mantener y mejorar la seguridad y el bienestar en nuestras ciudades.

Les deseo éxito en este emocionante viaje y espero ver los frutos de su arduo trabajo reflejados en nuestras comunidades.

# Índice

<b>1</b>	<b>Análisis del Caso de la Peste en Londres por John Snow: Aplicaciones en la Analítica Delictual</b>	<b>6</b>
1.1	Recolección de Datos . . . . .	6
1.2	Análisis Espacial . . . . .	6
1.3	Desafío a las Percepciones Previas . . . . .	6
1.4	Informando la Intervención . . . . .	7
1.5	Comunicación Visual . . . . .	7
<b>2</b>	<b>Análisis del Dilema de la Blindaje de Aviones durante la Segunda Guerra Mundial.</b>	<b>7</b>
2.1	Recolección de Datos . . . . .	8
2.2	Interpretación de Datos: . . . . .	9
2.3	Innovación Metodológica: . . . . .	9
2.4	Informando la Intervención: . . . . .	9
2.5	Comunicación de Hallazgos: . . . . .	9
<b>3</b>	<b>Teoría de las "Ventanas Rotas"</b>	<b>9</b>
3.1	Experimento de Zimbardo . . . . .	10
<b>4</b>	<b>Criminología Ambiental</b>	<b>10</b>
4.1	Análisis Espacial . . . . .	11
4.2	Diseño Urbano y Arquitectura . . . . .	11
4.3	Rutinas y Actividades Diarias . . . . .	12
4.4	Supervisión y Control Social . . . . .	12
4.5	Análisis de Riesgo y Evaluación . . . . .	12
4.6	Tecnologías de Información Geográfica (GIS) y Modelización . . . . .	12
<b>5</b>	<b>Teoría de "Puntos Calientes"</b>	<b>12</b>
5.1	Factores de importancia . . . . .	12
<b>6</b>	<b>Teoría del "Cisne Negro"</b>	<b>13</b>
6.1	Características de los Cisnes Negros . . . . .	13
6.2	Impacto en la Sociedad . . . . .	14
6.3	Implicancias para la Toma de Decisiones . . . . .	14
6.4	Crítica a los Métodos Predictivos Tradicionales . . . . .	14
6.5	Concepto de Antifragilidad . . . . .	14
<b>7</b>	<b>Excel Funciones Básicas y Avanzadas</b>	<b>16</b>
7.1	Historia de Excel: Multiplan . . . . .	16
7.2	Nacimiento de Excel . . . . .	16
7.3	Expansión a Windows y Crecimiento . . . . .	16
7.4	Datos Técnicos de Excel . . . . .	16
7.5	Características principales de Excel . . . . .	17
7.6	Importancia de Excel en el Análisis Delictual . . . . .	17
7.7	Funciones Aritméticas en Excel . . . . .	18
7.8	Funciones de texto . . . . .	18
7.9	Funciones de Fecha y Hora . . . . .	18
7.10	Funciones Lógicas . . . . .	19
7.11	Funciones de búsqueda . . . . .	19

7.12	Funciones Avanzadas de Análisis de Datos en Excel . . . . .	19
7.13	Tablas Dinámicas . . . . .	19
7.14	Funciones de Estadística . . . . .	20
7.15	Importación de Datos Delictuales . . . . .	20
7.16	Análisis Espacial . . . . .	20
7.17	Análisis Temporal . . . . .	20
7.18	Visualización de Datos . . . . .	20
<b>8</b>	<b>Análisis Exploratorio de Datos en Excel</b>	<b>21</b>
8.1	Medidas de Tendencia Central . . . . .	21
8.1.1	Media . . . . .	21
8.1.2	Mediana . . . . .	21
8.1.3	Moda . . . . .	21
8.2	Medidas de Dispersión . . . . .	21
8.2.1	Desviación Estándar . . . . .	21
8.3	Clasificación . . . . .	22
<b>9</b>	<b>Gráficos en Excel para Análisis Delictual</b>	<b>22</b>
9.1	Selección de Datos: . . . . .	22
9.2	Tipos de Gráficos Básicos y su Utilidad: . . . . .	22
9.3	Gráficos Avanzados y su Utilidad: . . . . .	22
9.4	Recomendaciones de Colores: . . . . .	23
9.5	Personalización del Gráfico: . . . . .	23
9.6	Análisis: . . . . .	23
9.7	Interpretación y Comunicación: . . . . .	23
9.8	Guardar y Exportar: . . . . .	24
<b>10</b>	<b>Correlación versus Causalidad en Análisis Criminal</b>	<b>26</b>
<b>11</b>	<b>Correlación de Pearson en Análisis Criminal</b>	<b>26</b>
<b>12</b>	<b>Regresión Lineal en Análisis Criminal</b>	<b>29</b>
<b>13</b>	<b>Proceso de identificación de Patrones Delictuales en Excel</b>	<b>32</b>
13.1	Recolección y Preparación de Datos: . . . . .	32
13.2	Análisis Exploratorio de Datos (EDA): . . . . .	33
13.3	Visualización de Datos: . . . . .	33
13.4	Análisis Temporal y Espacial: . . . . .	33
13.5	Búsqueda y Análisis de Patrones Específicos: . . . . .	33
13.6	Correlación y Regresión: . . . . .	33
13.7	Interpretación y Comunicación de Resultados: . . . . .	34
13.8	Recomendaciones: . . . . .	35

# Capítulo 1

## Antecedentes Históricos y fundamentos del Análisis de Datos

# 1. Análisis del Caso de la Peste en Londres por John Snow: Aplicaciones en la Analítica Delictual

El caso de la peste en Londres analizado por John Snow representa un hito histórico en el ámbito de la epidemiología y del análisis de datos geospaciales. Aunque se refiere a un evento de salud pública, ofrece lecciones valiosas aplicables al análisis delictual [1, 2].

## Antecedentes

En 1854, Londres fue azotada por una severa epidemia de cólera. John Snow, un médico inglés, mediante la observación meticulosa y el análisis de datos, pudo identificar una bomba de agua ubicada en Broad Street como el epicentro de la epidemia. Snow recopiló datos sobre los lugares de residencia de las personas afectadas y los ubicó en un mapa de la ciudad. Este mapa mostraba una clara concentración de casos alrededor de la bomba de Broad Street [3, 2].



Figura 1: John Snow, médico higienista inglés, considerado uno de los padres de la epidemiología, se hizo conocido por la solución que dio a un brote de cólera en el barrio de Soho, en Inglaterra, partiendo de los principios básicos del análisis espacial.

### 1.1. Recolección de Datos

Snow recolectó datos meticulosos sobre los casos de cólera, incluyendo las ubicaciones de las residencias de los afectados y las bombas de agua que utilizaban. De manera similar, los analistas delictuales deben recolectar y organizar datos precisos sobre incidentes delictuales y otras variables relevantes.

### 1.2. Análisis Espacial

Utilizando un mapa de la ciudad, Snow fue capaz de visualizar la concentración de casos alrededor de la bomba de agua en Broad Street, lo que le permitió identificar una clasificación espacial [4]. Los analistas delictuales utilizan enfoques similares, como el análisis de puntos calientes, para identificar áreas de alta actividad delictiva.

### 1.3. Desafío a las Percepciones Previas

Snow desafió la teoría dominante de la época que sugería que el cólera se propagaba a través del aire. Su análisis proporcionó evidencia de que, en realidad, se propagaba a través del agua

contaminada [2]. Los analistas delictuales también pueden descubrir, a través del análisis de datos, que ciertas percepciones o suposiciones sobre el crimen y la delincuencia pueden ser incorrectas.

#### **1.4. Informando la Intervención**

Gracias al análisis de Snow, las autoridades pudieron tomar medidas específicas para contener la epidemia, como desactivar la bomba de agua en Broad Street [1]. En el contexto delictual, el análisis de datos puede informar estrategias de intervención más efectivas, como el despliegue de recursos en áreas de alta criminalidad.

#### **1.5. Comunicación Visual**

Snow utilizó un mapa para comunicar sus hallazgos, una herramienta poderosa para mostrar visualmente la clasificación entre la ubicación de las bombas de agua y los casos de cólera [3]. Los analistas delictuales también utilizan mapas y otras herramientas de visualización para comunicar patrones delictuales y análisis a otras partes interesadas.

### **Conclusión del caso**

Este caso ilustra cómo el análisis de datos geoespaciales puede ser crucial para identificar patrones y fuentes de problemas graves. En el contexto delictual, los analistas pueden emplear métodos similares para identificar, por ejemplo, áreas de alta criminalidad y comprender mejor los patrones espaciales y temporales del crimen.

## **2. Análisis del Dilema de la Blindaje de Aviones durante la Segunda Guerra Mundial.**

Durante la Segunda Guerra Mundial, el dilema del blindaje de aviones tratados por Abraham Wald, famoso estadístico de la época, se aplicó a una variedad de aeronaves militares utilizadas por los Aliados. Los aviones involucrados eran principalmente bombarderos y cazadores, ya que eran los que enfrentaban mayores riesgos durante las misiones de combate. Algunos ejemplos de aviones de esa época incluyen el B-17 Flying Fortress, el B-24 Liberator y el P-51 Mustang, entre otros.

La meta era determinar cómo y dónde aplicar blindaje adicional a estas aeronaves para maximizar su supervivencia en el campo de batalla sin comprometer su rendimiento debido al peso adicional. Wald y su equipo del Statistical Research Group (SRG) de la Universidad de Columbia se embarcaron en un análisis detallado para abordar este problema, y sus conclusiones ayudaron a guiar las decisiones sobre cómo blindar los aviones de manera efectiva. La relevancia histórica del trabajo de Wald sigue siendo un ejemplo clásico de cómo el análisis estadístico cuidadoso y la interpretación correcta de los datos pueden informar decisiones críticas en situaciones de vida o muerte.

### **Antecedentes**

Durante la Segunda Guerra Mundial, los responsables en las pistas de vuelo llevaban a cabo un seguimiento de los aviones que retornaban de las misiones con daños causados por fuego de metralla. Estos daños se localizaban principalmente en las alas, el fuselaje y la cola. La sugerencia inicial fue reforzar estas áreas. Sin embargo, Abraham Wald argumentó que esta interpretación de los datos era engañosa, ya que se centraba únicamente en los aviones que habían regresado con éxito y no en

aquellos que habían sido derribados.

Wald sugirió que las áreas que no presentaban daños en los aviones que regresaban (como los motores y las áreas cercanas a la cabina) eran, en realidad, las zonas críticas que necesitaban blindaje adicional. Su razonamiento era que los aviones que recibían daños en estas áreas críticas probablemente eran derribados y, por lo tanto, no regresaban. Esto implicaba que las áreas sin daño en los aviones que regresaban eran precisamente las áreas que deberían ser blindadas para mejorar la supervivencia de los aviones en futuras misiones. Esta perspectiva contraintuitiva de Wald proporcionó una valiosa insight sobre cómo interpretar los datos de manera adecuada para informar decisiones cruciales, como dónde aplicar blindaje adicional en los aviones de combate para maximizar su supervivencia en el campo de batalla.

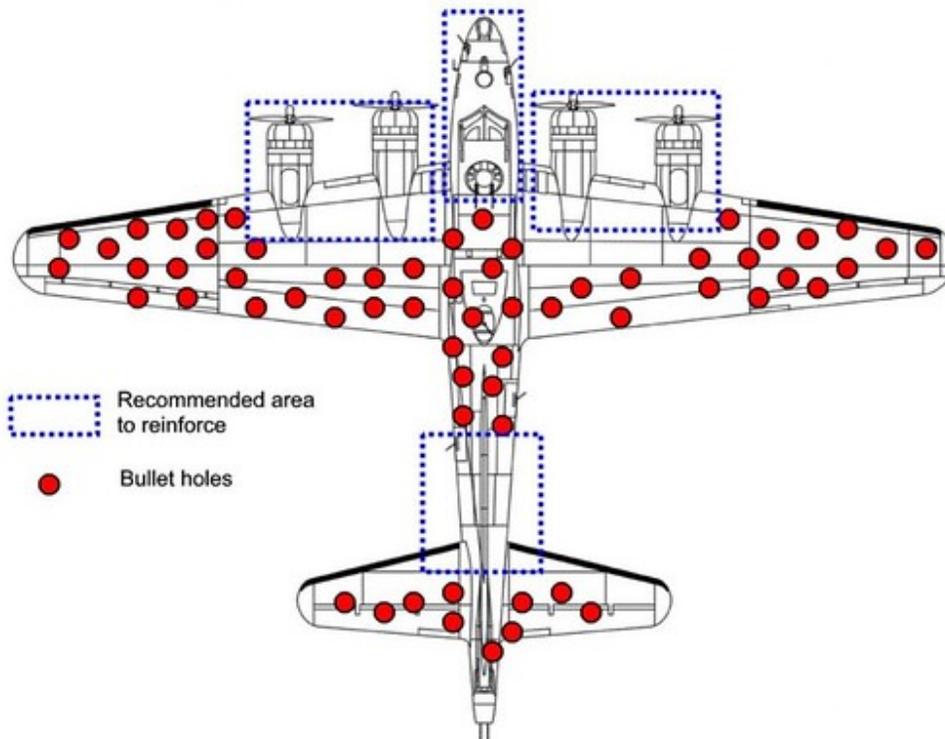


Figura 2: Los puntos rojos solo representaban el daño en los aviones que pudieron regresar sin caer derribados. Las áreas que realmente deberían reforzar, eran los lugares donde no había impactos, porque esos son los lugares donde el avión no sobreviviría al ser golpeado. Este fenómeno se llama sesgo de supervivencia. Es cuando miramos las cosas que sobrevivieron cuando en realidad deberíamos centrarnos en las que no están a nuestro alcance.

## 2.1. Recolección de Datos

Wald basó su análisis en los datos disponibles, aunque limitados, con estos pudo sacar conclusiones acertadas sobre los daños observados en los aviones que regresaban de las misiones de combate. De manera análoga, los analistas delictuales deben recolectar y examinar datos sobre incidentes delictuales, aun cuando estos datos parezcan inicialmente incompletos o sesgados.

## **2.2. Interpretación de Datos:**

Wald enfatizó la importancia de comprender y ajustar los sesgos inherentes en los datos antes de derivar conclusiones. Este enfoque es igualmente crítico para los analistas delictuales, quienes deben estar atentos a los sesgos y asegurarse de interpretar los datos delictuales de manera precisa.

## **2.3. Innovación Metodológica:**

Mediante su enfoque innovador, Wald pudo ofrecer una solución contraintuitiva que resultó ser efectiva. Los analistas delictuales, por su parte, deben estar dispuestos a aplicar métodos analíticos innovadores para adquirir nuevas perspectivas sobre la delincuencia y la seguridad.

## **2.4. Informando la Intervención:**

La interpretación adecuada de los datos por parte de Wald condujo a intervenciones efectivas para el blindaje de los aviones de manera más eficaz. De manera similar, un análisis delictual preciso puede guiar estrategias de intervención eficaces para abordar la delincuencia y mejorar la seguridad.

## **2.5. Comunicación de Hallazgos:**

Wald comunicó con éxito sus hallazgos, lo que permitió la implementación de sus recomendaciones. De igual manera, los analistas delictuales deben ser capaces de comunicar sus hallazgos de manera efectiva a las partes interesadas para influir en la toma de decisiones.

Este caso ha sido ampliamente citado en la literatura sobre estadísticas, análisis de decisiones y otras áreas relacionadas. Una fuente recomendada para explorar más sobre este caso es el libro "How Not to Be Wrong: The Power of Mathematical Thinking" de Jordan Ellenberg, donde se discute el Dilema del Blindaje de Aviones en el contexto de la interpretación correcta de los datos y la toma de decisiones informada [5].

## **3. Teoría de las "Ventanas Rotas"**

La teoría de las "Ventanas Rotas" fue presentada por James Q. Wilson y George L. Kelling en un artículo publicado en 1982 en la revista "Atlantic Monthly" bajo el título "Broken Windows: The police and neighborhood safety" [6]. Esta teoría plantea que mantener y monitorear los entornos urbanos para evitar pequeños crímenes como el vandalismo, el desorden público y el abandono puede ayudar a crear una atmósfera de orden y legalidad, previniendo así crímenes más graves.

Según Wilson y Kelling, una ventana rota en un edificio es una invitación para que se cometan más actos de vandalismo. Si no se repara, puede llevar a una escalada de desorden y eventualmente a crímenes más serios. La idea es que el desorden y la incivilidad en una comunidad envían una señal de que a nadie le importa, por lo que el control social se rompe. La teoría sugiere que la policía debería enfocarse en mantener el orden público y prevenir pequeñas ofensas para crear un ambiente de orden y legalidad.

El artículo provocó un cambio en las políticas de aplicación de la ley, con muchas ciudades adoptando el enfoque de "tolerancia cero" hacia el desorden y las ofensas menores, aunque la efectividad de estas políticas y la interpretación de la teoría han sido temas de debate en las décadas siguientes.

### 3.1. Experimento de Zimbardo

En el experimento de Zimbardo, se dejaron dos autos abandonados en dos lugares muy diferentes: uno en el Bronx, Nueva York, y el otro en Palo Alto, California. El auto en el Bronx fue vandalizado casi inmediatamente, mientras que el auto en Palo Alto permaneció intacto. Sin embargo, después de que Zimbardo mismo dañó deliberadamente el auto en Palo Alto con un martillo, los transeúntes comenzaron a vandalizarlo también.



Figura 3: La Teoría de las Ventanas Rotas sugiere que los signos visibles de desorden y negligencia, como ventanas rotas, pintadas y basura, pueden crear un entorno que fomente más el vandalismo y el comportamiento antisocial.

El experimento de Zimbardo mostró cómo la presencia de desorden (en este caso, un auto ya vandalizado) puede conducir a más desorden. Aunque este experimento no fue directamente relacionado con la Teoría de las Ventanas Rotas en ese momento, proporcionó una base experimental para las ideas que Wilson y Kelling propondrían más tarde.

La Teoría de las Ventanas Rotas fue llevada más allá en los años 80 y 90, especialmente en la Ciudad de Nueva York, donde la estrategia de "tolerancia cero" para el desorden y los crímenes menores fue implementada por el comisionado de policía William Bratton. Se argumentó que esta estrategia condujo a una reducción en los crímenes mayores, aunque esta interpretación y la eficacia de la política han sido ampliamente debatidas.

Este experimento y la teoría subyacente proporcionan una perspectiva fascinante sobre cómo el desorden y la percepción de desorden en un entorno pueden afectar el comportamiento de las personas y, potencialmente, llevar a una mayor delincuencia y desorden.

## 4. Criminología Ambiental

El libro *Environmental Criminology* por Paul J. Brantingham y Patricia L. Brantingham (1981) aborda la relación entre el ambiente, la geografía, y el crimen, proponiendo un enfoque novedoso

para entender cómo y por qué ocurren los crímenes en ciertos lugares [7]. Los autores introducen el concepto de criminología ambiental, que busca entender cómo los factores espaciales y ambientales influyen en la actividad delictiva.

Se destaca la importancia de considerar el entorno físico y el diseño urbano en la prevención del crimen. Los autores discuten cómo ciertos lugares o "puntos calientes" pueden ser propensos a la actividad criminal debido a su diseño o a la falta de supervisión. También se explora cómo las rutinas diarias de las personas interactúan con el ambiente para influir en las oportunidades de crimen.

El análisis espacial y la modelización geográfica son presentados como herramientas valiosas para identificar patrones de crimen y desarrollar estrategias de prevención. Este libro proporciona un fundamento teórico y metodológico para los estudios posteriores en criminología ambiental, destacando cómo el análisis del ambiente y del espacio puede contribuir a una comprensión más profunda y a la prevención del crimen.

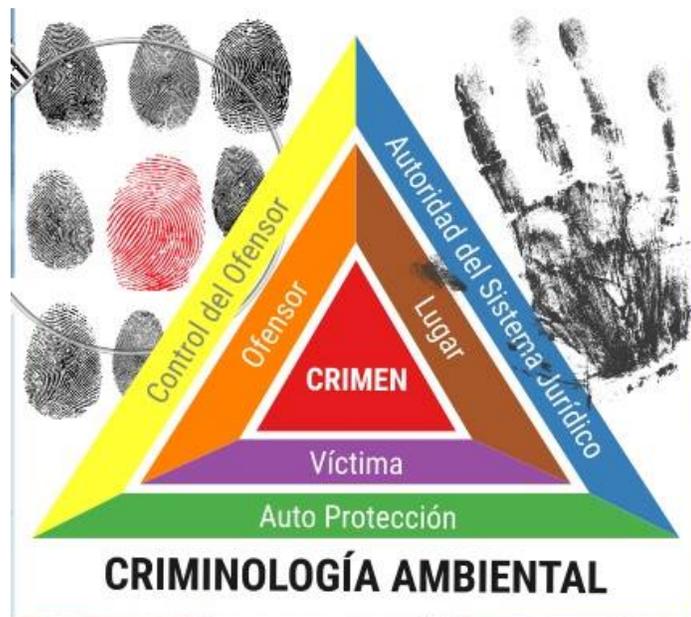


Figura 4: Es una rama de la criminología que estudia el espacio o lugar y el momento en el que se ha cometido algún hecho delictivo. Son diferentes perspectivas teóricas las que intentan explicar y prevenir el delito a partir de las influencias situacionales: el diseño de las calles, de los edificios, la luminosidad, la cuantía de árboles, de garajes, de personas, etc. Todo ello afecta de una manera o de otra a la comisión o no de un acto criminal.

#### 4.1. Análisis Espacial

- Identificación de "puntos calientes" del crimen o áreas que tienen una alta concentración de actividad delictiva.
- Análisis de la distribución geográfica del crimen y cómo ciertas características del entorno pueden influir en esta distribución.

#### 4.2. Diseño Urbano y Arquitectura

- La influencia del diseño urbano, como la iluminación, la visibilidad y el acceso, en la prevención del crimen.

- Cómo la configuración de un espacio puede afectar las oportunidades para el crimen.

#### **4.3. Rutinas y Actividades Diarias**

- Cómo las rutinas diarias de las personas interactúan con el entorno y pueden crear oportunidades para el crimen.
- La relación entre la movilidad humana y la exposición al riesgo de victimización.

#### **4.4. Supervisión y Control Social**

- Cómo la presencia de supervisión, como la vigilancia policial o la supervisión comunitaria, puede desincentivar la actividad delictiva.
- El papel de la comunidad en la prevención del crimen y el mantenimiento del orden.

#### **4.5. Análisis de Riesgo y Evaluación**

- Evaluación del riesgo de delito en diferentes áreas y cómo la información sobre el riesgo puede ser utilizada para informar estrategias de prevención.

#### **4.6. Tecnologías de Información Geográfica (GIS) y Modelización**

- Utilización de tecnologías GIS para analizar datos de crimen espacialmente.
- Modelos matemáticos y estadísticos para predecir patrones de crimen y evaluar intervenciones.

Estos aspectos resaltan cómo la interacción entre los individuos y el entorno puede influir en la ocurrencia del crimen y cómo la comprensión de estos factores puede ser utilizada para desarrollar estrategias efectivas de prevención del crimen.

### **5. Teoría de "Puntos Calientes"**

La teoría de los *puntos calientes* del crimen (*Hot Spots of Crime*) propuesta por Sherman et al. [8], postula que el crimen tiende a estar concentrado en áreas geográficas específicas en lugar de estar distribuido de manera uniforme. Los "puntos calientes" son áreas que experimentan una incidencia del crimen significativamente más alta en comparación con otras áreas. La identificación de estos puntos calientes puede ayudar a las autoridades y a los responsables de la formulación de políticas a dirigir recursos de manera más efectiva para la prevención del crimen y la intervención.

#### **5.1. Factores de importancia**

Los autores exploran cómo las actividades rutinarias y las características del entorno en estos puntos calientes contribuyen a la prevalencia del crimen. Argumentan que la estructura física, la densidad de población, la iluminación, la vigilancia y otras características ambientales pueden influir en la oportunidad para que ocurran crímenes. Por ejemplo, un área con poca iluminación y escasa vigilancia podría proporcionar un entorno propicio para actividades delictivas.

Además, las actividades rutinarias de las personas, como dónde trabajan, dónde socializan y cómo se transportan, pueden interactuar con estas características ambientales para crear oportunidades



Figura 5: Concentración de crímenes en áreas específicas denominadas "puntos calientes".

para el crimen. Por lo tanto, los patrones de vida diarios de las personas pueden contribuir a la formación de puntos calientes del crimen.

Este análisis basado en la ubicación ofrece una comprensión profunda de cómo y por qué ciertas áreas se vuelven propensas al crimen, lo que a su vez puede informar estrategias de prevención del crimen basadas en la evidencia. Por ejemplo, mejorar la iluminación o aumentar la vigilancia en un punto caliente identificado podría reducir las oportunidades para el crimen y, por lo tanto, la incidencia del crimen en esa área.

El análisis de los puntos calientes del crimen se ha convertido en una herramienta esencial en la criminología moderna y la justicia penal, proporcionando un marco para el despliegue estratégico de recursos y la planificación de intervenciones para abordar el crimen de manera más efectiva.

## 6. Teoría del "Cisne Negro"

El término Cisne Negro proviene de una antigua creencia que se remonta a la Roma antigua, donde se asumía que todos los cisnes eran blancos, ya que eso era lo único que se había observado. Esta suposición perduró hasta el siglo XVII, cuando los exploradores europeos descubrieron cisnes negros en Australia. El descubrimiento desafió una creencia largamente sostenida y sirvió como un poderoso recordatorio de las limitaciones del conocimiento humano y la expectativa.

El concepto del Cisne Negro fue popularizado por Nassim Nicholas Taleb en su libro "The Black Swan: The Impact of the Highly Improbable" [9]. Taleb argumenta que existe una clase de eventos, denominados Cisnes Negros, que son altamente improbables, tienen un impacto extremadamente grande y son difíciles de predecir, pero son más previsibles retrospectivamente.

### 6.1. Características de los Cisnes Negros

Según Taleb, los Cisnes Negros poseen tres características principales:

1. **Raros:** Son eventos que se encuentran fuera del reino de las expectativas regulares porque nada en el pasado puede apuntar convincentemente a su posibilidad.
2. **Impacto extremo:** Su ocurrencia tiene un efecto significativo y generalizado.



Figura 6: El colapso financiero de 2008 o crisis Subprime, es un ejemplo de un Cisne Negro.<sup>en</sup> el sentido de que fue un evento altamente improbable (según las percepciones previas), tuvo un impacto extremadamente grande, y solo se hizo "predecible.<sup>o</sup> comprensible después de lo ocurrido.

3. **Previsibilidad retrospectiva:** Después de su ocurrencia, se explican a posteriori como si fueran predecibles.

## 6.2. Impacto en la Sociedad

Taleb sugiere que los Cisnes Negros son a menudo consecuencia de la falta de conocimiento y comprensión del mundo, y de la tendencia humana a buscar explicaciones simples para eventos complejos. Estos eventos pueden tener un impacto profundo en la economía, la tecnología, la política, la ciencia, entre otros campos, alterando el status quo y desafiando las percepciones existentes.

## 6.3. Implicancias para la Toma de Decisiones

Taleb aboga por una revisión de cómo se abordan los riesgos y se toman decisiones en presencia de incertidumbre. Aconseja la construcción de sociedades y sistemas más robustos que puedan resistir shocks inesperados, en lugar de intentar predecir lo impredecible.

## 6.4. Crítica a los Métodos Predictivos Tradicionales

Taleb critica duramente los métodos estadísticos tradicionales que, según él, a menudo fallan en la captura de la probabilidad y el impacto de estos eventos de Cisne Negro. Argumenta que los modelos que ignoran o minimizan la posibilidad de Cisnes Negros están fundamentalmente incompletos y pueden llevar a decisiones erróneas y a una falsa sensación de seguridad.

## 6.5. Concepto de Antifragilidad

En trabajos posteriores, Taleb expande su teoría introduciendo el concepto de antifragilidad, que describe sistemas que se benefician de la volatilidad y la incertidumbre, en contraposición a la mera resistencia o robustez frente a los Cisnes Negros.

Este marco conceptual propuesto por Taleb ha generado amplias discusiones y reflexiones sobre cómo se entiende y se gestiona la incertidumbre en múltiples campos, enfatizando la importancia de prepararse para lo desconocido en lugar de simplemente confiar en predicciones basadas en datos históricos.

## Capítulo 2

# Herramientas Esenciales de Excel para el Análisis de Datos

## 7. Excel Funciones Básicas y Avanzadas

Microsoft Excel es una herramienta de hoja de cálculo que forma parte del paquete de software de Microsoft Office. Es ampliamente utilizado en diferentes campos debido a su capacidad para realizar cálculos complejos, análisis de datos, gráficos y la automatización de tareas a través de macros. Excel proporciona un espacio de trabajo basado en celdas donde los usuarios pueden ingresar datos, fórmulas y utilizar una amplia variedad de funciones predefinidas para manipular y analizar información.

### 7.1. Historia de Excel: Multiplan

En el año 1982, Microsoft incursionó en el mercado de las hojas de cálculo con el lanzamiento de Multiplan, un programa diseñado para competir con VisiCalc de VisiCorp, el cual era el primer software de hoja de cálculo para computadoras personales en ese momento. A pesar de la versatilidad de Multiplan, que podía ejecutarse en varias computadoras, no logró superar la popularidad y ventas de Lotus 1-2-3, una hoja de cálculo que se estaba consolidando rápidamente como una de las más preferidas durante la década de 1980.

### 7.2. Nacimiento de Excel

Viendo la necesidad de innovar y ofrecer un producto superior, Microsoft decidió reinventar su oferta de hoja de cálculo. En 1985, dejaron de lado Multiplan y lanzaron la primera versión de Microsoft Excel. Sin embargo, esta versión inicial solo estaba disponible para las computadoras Macintosh. Aunque fue un comienzo modesto, sentó las bases para lo que vendría después.

### 7.3. Expansión a Windows y Crecimiento

Dos años más tarde, en 1987, Microsoft lanzó Excel 2.0, una versión renovada y mejorada que ahora era compatible con la plataforma Windows. Este fue un punto de inflexión crucial en la historia de Excel, ya que la compatibilidad con Windows abrió las puertas a una base de usuarios mucho más amplia. Desde este punto, Excel comenzó a ganar terreno y su popularidad creció exponencialmente. Con el paso de los años, Microsoft continuó mejorando y expandiendo las capacidades de Excel, incorporando una vasta gama de funciones, herramientas de análisis y visualización de datos, y opciones de automatización a través de la programación en VBA (Visual Basic for Applications). Estas innovaciones fueron cruciales para consolidar a Excel como la hoja de cálculo más utilizada en el mundo.

### 7.4. Datos Técnicos de Excel

- **Lenguajes de Programación:** C++, Objective-C
- **Desarrollador:** Microsoft
- **Lanzamiento Inicial:** 30 de septiembre de 1985
- **Licencia:** Software Propietario
- **Tipo de Programa:** Hoja de Cálculo
- **Última Versión Estable:** 2021 (5 de octubre de 2021)



Figura 7: El uso eficiente de Excel puede acelerar el proceso de análisis delictivo, permitiendo a los analistas concentrarse en la interpretación y comunicación de los hallazgos, y contribuyendo de manera significativa en la toma de decisiones informadas en el ámbito de la seguridad pública y privada.

Excel ha continuado evolucionando a lo largo de los años, con Microsoft lanzando nuevas versiones que incluyen una gama aún más amplia de características y funcionalidades. Su papel ha sido crucial en muchos campos, incluyendo el análisis criminal, donde ofrece a los profesionales las herramientas necesarias para recopilar, analizar y visualizar datos de manera efectiva.

### 7.5. Características principales de Excel

- **Hojas de Cálculo:** Excel permite a los usuarios trabajar con hojas de cálculo donde los datos se pueden ingresar en una estructura de tabla con filas y columnas.
- **Fórmulas y funciones:** Excel tiene una vasta colección de fórmulas y funciones que permiten a los usuarios realizar cálculos desde básicos hasta avanzados.
- **Gráficos y Visualización de Datos:** Los usuarios pueden crear varios tipos de gráficos para visualizar datos y obtener insights.
- **Tablas Dinámicas:** Esta poderosa característica permite a los usuarios analizar grandes conjuntos de datos y resumir información de manera efectiva.
- **Automatización y Macros:** A través del uso de VBA (Visual Basic para Aplicaciones), los usuarios pueden crear macros para automatizar tareas repetitivas.
- **Análisis de Datos y Herramientas de Solver:** Excel proporciona herramientas para análisis estadístico, análisis de tendencias y optimización.

### 7.6. Importancia de Excel en el Análisis Delictual

El análisis criminal es un campo que requiere la recopilación, manejo y análisis de grandes cantidades de datos. Excel, con sus capacidades de manipulación de datos, análisis y visualización, se convierte en una herramienta invaluable para los analistas criminales. Las características mencionadas anteriormente, como las tablas dinámicas y las funciones de análisis de datos, pueden ser especialmente útiles para examinar tendencias delictuales, identificar patrones y comunicar hallazgos de manera efectiva a las partes interesadas.

En el análisis criminal, Excel puede ser utilizado para:

- **Recopilación y Organización de Datos:** Excel permite a los analistas criminales recopilar y organizar datos delictuales de manera estructurada y sistemática.

- **Análisis Estadístico:** Las funciones estadísticas de Excel pueden ayudar a los analistas a entender las tendencias delictuales y a realizar comparaciones entre diferentes conjuntos de datos.
- **Visualización de Datos criminales:** Los analistas pueden utilizar las herramientas de gráficos de Excel para visualizar datos delictuales y presentar información de manera clara y comprensible.
- **Predicción y Modelización:** Excel proporciona herramientas que pueden ayudar a los analistas a desarrollar modelos predictivos basados en datos históricos delictuales.
- **Automatización de Tareas Repetitivas:** Las macros en Excel pueden ayudar a automatizar tareas repetitivas, lo que ahorra tiempo y reduce la posibilidad de errores.

### 7.7. Funciones Aritméticas en Excel

Las funciones aritméticas son esenciales para realizar cálculos básicos en Excel. Algunas de las funciones más utilizadas incluyen:

- **SUMA:** Suma todos los números en un rango de celdas.
- **PROMEDIO:** Calcula el promedio de un grupo de números.
- **MEDIANA:** Encuentra el valor medio de un grupo de números.
- **MAX:** Devuelve el valor más grande de un conjunto.
- **MIN:** Devuelve el valor más pequeño de un conjunto.

### 7.8. Funciones de texto

Manipular texto es esencial, especialmente cuando se trabaja con datos que incluyen descripciones o identificadores. Algunas funciones útiles son:

- **IZQUIERDA:** Devuelve un número específico de caracteres desde el inicio de una cadena.
- **DERECHA:** Devuelve un número específico de caracteres desde el final de una cadena.
- **CONCATENAR:** Combina varias cadenas en una sola.
- **SUPERIOR:** Convierte texto en mayúsculas.
- **INFERIOR:** Convierte texto en minúsculas.

### 7.9. Funciones de Fecha y Hora

Las fechas y horas son fundamentales en el análisis criminal, especialmente para identificar tendencias temporales. Algunas funciones clave incluyen:

- **FECHA:** Devuelve una fecha basada en año, mes y día.
- **HORA:** Devuelve una hora basada en horas, minutos y segundos.
- **AHORA:** Devuelve la fecha y hora actuales.
- **DÍA:** Extrae el día de una fecha.
- **MES:** Extrae el mes de una fecha.
- **AÑO:** Extrae el año de una fecha.

## 7.10. Funciones Lógicas

Las funciones lógicas permiten realizar análisis condicionales, lo que es útil para categorizar o filtrar datos basados en ciertas condiciones. Algunas funciones lógicas populares son:

- **SI:** Realiza un test lógico y devuelve un valor si el test es verdadero y otro si es falso.
- **Y:** Devuelve VERDADERO si todas las condiciones son verdaderas.
- **O:** Devuelve VERDADERO si al menos una de las condiciones es verdadera.
- **NO:** Devuelve el valor lógico opuesto.

## 7.11. Funciones de búsqueda

Estas funciones son esenciales cuando se necesita buscar y recuperar información específica en grandes conjuntos de datos:

- **BUSCARV:** Busca un valor en la columna más a la izquierda y devuelve un valor en la misma fila desde una columna especificada.
- **BUSCARH:** Similar a BUSCARV pero busca en filas en lugar de columnas.
- **ÍNDICE:** Devuelve el valor de una celda en una tabla o rango.
- **COINCIDIR:** Busca un valor y devuelve la posición relativa de ese ítem dentro de un rango.

## 7.12. Funciones Avanzadas de Análisis de Datos en Excel

Para realizar análisis detallados en conjuntos de datos delicados, Excel proporciona funciones específicas que permiten contar, sumar o calcular el promedio basado en condiciones específicas. Algunas de estas funciones son:

- **CONTAR.SI:** Cuenta el número de celdas dentro de un rango que cumplen con una condición especificada.
- **SUMAR.SI:** Suma las celdas especificadas por una condición dada.
- **PROMEDIO.SI:** Calcula el promedio de las celdas que cumplen una condición específica.

## 7.13. Tablas Dinámicas

Las tablas dinámicas son herramientas poderosas que permiten explorar, analizar y sintetizar grandes conjuntos de datos delictuales de una manera más estructurada y significativa. Permiten:

- Resumir y analizar grandes conjuntos de datos.
- Crear reportes interactivos.
- Identificar patrones y tendencias.
- Organizar y comparar datos de manera efectiva.

#### 7.14. Funciones de Estadística

El análisis estadístico es crucial para entender y interpretar los datos delictuales. Excel proporciona una serie de funciones estadísticas como:

- **DESVEST:** Calcula la desviación estándar, una medida de la cantidad de variación o dispersión de un conjunto de valores.
- **VAR:** Calcula la varianza de una muestra.
- **CORREL:** Calcula el coeficiente de correlación entre dos conjuntos de datos, proporcionando una medida de la relación entre las variables.

#### 7.15. Importación de Datos Delictuales

La importación de datos delictuales es el primer paso crucial en el análisis criminológico. Excel permite la importación de bases de datos relacionadas con la delincuencia de diferentes fuentes y formatos como CSV, TXT, o incluso desde bases de datos externas. Una vez importados, estos datos pueden ser organizados, filtrados y preparados para el análisis posterior.

#### 7.16. Análisis Espacial

El análisis espacial se refiere al estudio de patrones y procesos geográficos. En el contexto del análisis criminal, permite investigar la distribución geográfica de los delitos. Aunque Excel no es una herramienta GIS (Sistema de Información Geográfica) completa, puede proporcionar funcionalidades básicas de geolocalización y análisis espacial, especialmente cuando se complementa con otros plugins o software especializado. La geolocalización de incidentes delictivos y la visualización de datos en mapas pueden ayudar a identificar áreas de alta criminalidad y a entender mejor los patrones delictivos espaciales.

#### 7.17. Análisis Temporal

El análisis temporal implica el estudio de las tendencias delictuales a lo largo del tiempo. Excel ofrece funciones y herramientas que facilitan el análisis de series temporales, como la creación de gráficos de líneas o barras que representen la evolución del crimen durante un período específico. Al analizar las tendencias temporales, los analistas criminales pueden identificar patrones, estacionalidades o anomalías que podrían ser indicativas de problemas emergentes o la efectividad de las intervenciones de seguridad.

#### 7.18. Visualización de Datos

La visualización es un aspecto esencial del análisis criminal ya que facilita la interpretación y comunicación de los hallazgos. Excel proporciona una amplia gama de herramientas de visualización como gráficos de barras, líneas, dispersión, y paneles de control, que pueden ser utilizadas para representar datos delictuales de manera efectiva. A través de una visualización adecuada, los analistas pueden presentar información crítica de una manera que sea fácilmente comprensible para las partes interesadas, lo que contribuye a la toma de decisiones informada en el ámbito de la seguridad pública.

## 8. Análisis Exploratorio de Datos en Excel

El Análisis Exploratorio de Datos (en inglés: Exploratory Data Analysis por sus siglas EDA) es una fase crucial en cualquier proceso de análisis criminal que permite entender la estructura y las características básicas de los datos antes de proceder a análisis más complejos. Excel ofrece una gama de herramientas y funciones que facilitan la realización de AED. A continuación, se detallan algunos de los aspectos fundamentales de la analítica descriptiva y exploratoria en Excel.

### 8.1. Medidas de Tendencia Central

#### 8.1.1. Media

La media es el valor promedio de un conjunto de datos. Se calcula sumando todos los valores y dividiendo por la cantidad de valores.

$$\text{Fórmula :Media} = \frac{1}{n} \sum_{i=1}^n x_i$$

**Ejemplo en Excel:** Suponiendo que los datos están en el rango A1:A10, la fórmula sería:

=PROMEDIO (A1 : A10)

#### 8.1.2. Mediana

La mediana es el valor que se encuentra en el medio de un conjunto de datos ordenados.

*Fórmula en Excel :*

=MEDIANA (A1 : A10)

#### 8.1.3. Moda

La moda es el valor que ocurre con mayor frecuencia en un conjunto de datos.

*Fórmula en Excel :*

=MODA.MULT (A1 : A10)

### 8.2. Medidas de Dispersión

#### 8.2.1. Desviación Estándar

La desviación estándar mide la cantidad de variación o dispersión de un conjunto de valores.

$$\text{Fórmula :}\sigma = \sqrt{\frac{1}{n-1} \sum (x_i - \bar{x})^2}$$

**Ejemplo en Excel:** Suponiendo que los datos están en el rango A1:A10, la fórmula sería:

=DESVEST (A1 : A10)

### 8.3. Clasificación

La clasificación implica ordenar los datos en un orden específico, ya sea ascendente o descendente. **Ejemplo en Excel:** Suponiendo que los datos están en el rango A1:A10, para ordenarlos de manera ascendente:

=ORDENAR (A1 : A10)

Este análisis exploratorio básico en Excel permite tener una primera comprensión de los datos delictuales y preparar el terreno para análisis más profundos y específicos.

## 9. Gráficos en Excel para Análisis Delictual

El análisis criminal se beneficia significativamente de la visualización gráfica, ya que permite a los analistas y stakeholders entender las tendencias y patrones del delito de manera más clara y efectiva. Microsoft Excel ofrece una amplia gama de opciones de gráficos que pueden ser de gran ayuda. A continuación, se describen varios tipos de gráficos y cómo pueden ser utilizados en el análisis criminal:

### 9.1. Selección de Datos:

- Primero, es necesario seleccionar los datos que desea visualizar. Esta selección puede incluir valores numéricos, fechas, categorías, entre otros.

### 9.2. Tipos de Gráficos Básicos y su Utilidad:

- **Gráfico de Barras y Columnas:** Estos gráficos son excelentes para comparar valores entre diferentes categorías. Por ejemplo, comparar el número de delitos cometidos en diferentes barrios o durante diferentes meses. No son adecuados para series temporales o datos continuos.
- **Gráfico de Líneas:** Ideal para mostrar tendencias a lo largo del tiempo, como la evolución del delito a lo largo de los años. No son adecuados para comparar categorías individuales que no tienen una relación temporal o secuencial.
- **Gráfico de Pastel:** Útil para mostrar proporciones o porcentajes, como la distribución de diferentes tipos de delitos en una zona determinada. No son adecuados cuando tienes muchas categorías o si las diferencias entre los valores son mínimas.

### 9.3. Gráficos Avanzados y su Utilidad:

- **Gráfico de Área:** Este gráfico es eficaz para mostrar la magnitud de los valores a lo largo del tiempo y es útil para representar la acumulación de incidentes delictivos. No son adecuados para datos categóricos sin una relación temporal.
- **Gráfico de Dispersión:** Útil para identificar relaciones o correlaciones entre variables, como la relación entre el tiempo y la frecuencia de ciertos tipos de delitos. No es adecuado para datos categóricos o comparaciones entre diferentes categorías.
- **Histograma:** Este gráfico es beneficioso para visualizar la distribución de frecuencias de un conjunto de datos, como la frecuencia de delitos en diferentes rangos de tiempo. No es adecuado para comparar diferentes grupos o categorías.

- **Gráfico de Radar:** Puede ser utilizado para comparar múltiples variables o categorías, útil para comparar diferentes tipos de delitos en diversas áreas geográficas. No es adecuado para mostrar tendencias a lo largo del tiempo.
- **Gráfico de Burbujas:** Este gráfico es útil para representar tres dimensiones de datos, como la ubicación, el tiempo y la magnitud de los delitos. No es adecuado para datos categóricos simples.
- **Mapas de Calor:** Aunque requiere el uso de complementos adicionales, los mapas de calor en Excel pueden ayudar a visualizar la densidad delictual en diferentes áreas geográficas. No son adecuados para representar datos temporales o categorías no geográficas.

#### 9.4. Recomendaciones de Colores:

- Utilice colores contrastantes para separar diferentes categorías o series de datos en el gráfico.
- Evite el uso de colores brillantes o distractivos que puedan desviar la atención de los datos.
- Para gráficos de líneas o de barras, considere utilizar un esquema de color secuencial o divergente para representar la variación en los datos.
- Asegúrese de que los colores elegidos sean accesibles para los lectores con discromatopsia (dificultad para distinguir colores).

#### 9.5. Personalización del Gráfico:

- **Título del Gráfico:** Añada un título descriptivo al gráfico para que los espectadores entiendan de qué se trata.
- **Etiquetas de Eje:** Es crucial etiquetar los ejes correctamente para proporcionar contexto sobre los datos visualizados.
- **Leyenda:** Si su gráfico compara múltiples categorías o series de datos, una leyenda ayudará a los espectadores a entender qué representa cada color o forma en el gráfico.
- **Formato de Eje:** Ajuste el formato de los ejes para asegurarse de que los valores sean legibles y comprendidos correctamente.

#### 9.6. Análisis:

- Una vez creado el gráfico, utilícelo para analizar tendencias, comparaciones y patrones en los datos delictuales. Por ejemplo, podría identificar meses de alta actividad delictual o áreas con altas tasas de ciertos tipos de delitos.

#### 9.7. Interpretación y Comunicación:

- Interprete los hallazgos visualizados en el gráfico y comunique estos hallazgos a las partes interesadas pertinentes para informar las decisiones y estrategias de prevención del delito.

## 9.8. Guardar y Exportar:

- Guarde su trabajo y considere la posibilidad de exportar el gráfico como una imagen o PDF si necesita compartirlo fuera de Excel o incluirlo en informes.

Estos pasos básicos le ayudarán a comenzar con la visualización de datos delictuales en Excel. Con práctica, podrá explorar más opciones y personalizaciones para hacer gráficos más informativos y visualmente atractivos.

## Capítulo 3

# Análisis Estadístico y Patrones Delictuales Utilizando Excel

## 10. Correlación versus Causalidad en Análisis Criminal

La correlación y la causalidad son conceptos fundamentales en estadística y análisis criminal que, aunque relacionados, representan fenómenos diferentes.

**Correlación:** Este término se refiere a la relación estadística entre dos variables, donde cambios en una variable están asociados con cambios en otra. En el análisis criminal, por ejemplo, podría existir una correlación entre el aumento de la temperatura y la tasa de delitos violentos.

**Causalidad:** Implica una relación donde un evento o variable provoca o causa un efecto en otra variable. Por ejemplo, la implementación de un programa de vigilancia de Vigilantes privados podría causar una reducción en la tasa de delitos en un área específica.



Figura 8: La frase "correlación no implica causalidad" es un principio fundamental en estadística y ciencia de datos que indica que solo porque dos variables están correlacionadas, no significa que una cause la otra. La frase se ha utilizado ampliamente y su origen exacto es difícil de determinar.

## 11. Correlación de Pearson en Análisis Criminal

**Correlación de Pearson:** La Correlación de Pearson es una estadística que mide la fuerza y la dirección de la asociación entre dos variables continuas. El valor de la correlación varía entre -1 y 1, donde 1 indica una correlación positiva perfecta, -1 una correlación negativa perfecta, y 0 ninguna correlación.

Es vital en análisis criminal para entender, por ejemplo, cómo diferentes variables como el desempleo o la educación pueden estar correlacionadas con las tasas de delincuencia.

La fórmula para calcular la correlación de Pearson ( $r$ ) entre dos variables  $x$  y  $y$  es:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Desglosemos esta fórmula en pasos para entenderla mejor:

### 1. Cálculo de las medias:

- Calcular la media ( $\bar{x}$ ) de la variable  $x$  sumando todos los valores de  $x$  y dividiendo por el número total de observaciones.
- Calcular la media ( $\bar{y}$ ) de la variable  $y$  sumando todos los valores de  $y$  y dividiendo por el número total de observaciones.

$$\bar{x} = \frac{\sum x_i}{n}$$

$$\bar{y} = \frac{\sum y_i}{n}$$

### 2. Cálculo de las diferencias:

- Calcular la diferencia entre cada valor de  $x$  y la media de  $x$ , y la diferencia entre cada valor de  $y$  y la media de  $y$ .

$$d_{x_i} = x_i - \bar{x}$$

$$d_{y_i} = y_i - \bar{y}$$

### 3. Cálculo del numerador:

- Multiplicar cada pareja de diferencias obtenidas en el paso anterior y sumar todos estos productos.

$$\sum d_{x_i} \cdot d_{y_i}$$

### 4. Cálculo del denominador:

- Calcular la suma de las diferencias al cuadrado para cada variable y luego multiplicar estas sumas.

$$\sqrt{(\sum d_{x_i}^2) \cdot (\sum d_{y_i}^2)}$$

### 5. Cálculo de la Correlación de Pearson:

- Dividir el numerador por el denominador.

$$r = \frac{\sum d_{x_i} \cdot d_{y_i}}{\sqrt{(\sum d_{x_i}^2) \cdot (\sum d_{y_i}^2)}}$$

Esta fórmula proporciona el coeficiente de r que puede variar de -1 a 1, donde el signo indica la dirección de la correlación y el valor numérico, la magnitud de la correlación. En este contexto se resumen algunos criterios de interpretación:

- $-1,00$  = existe una correlación positiva perfecta. El índice indica una dependencia total entre las dos variables denominada relación directa: cuando una de ellas aumenta, la otra también lo hace en proporción constante.
- $-0,90$  = Correlación negativa muy fuerte
- $-0,75$  = Correlación negativa considerable
- $-0,50$  = Correlación negativa media
- $-0,10$  = Correlación negativa débil
- $0,00$  = entonces no existe relación lineal pero esto no necesariamente implica que las variables son independientes: pueden existir todavía relaciones no lineales entre las dos variables.
- $0,10$  = Correlación positiva débil
- $0,50$  = Correlación positiva media
- $0,75$  = Correlación positiva considerable
- $0,90$  = Correlación positiva muy fuerte
- $1,00$  = existe una correlación negativa perfecta. El índice indica una dependencia total entre las dos variables llamada relación opuesta: cuando una de ellas aumenta, la otra cambia su signo en proporción constante.

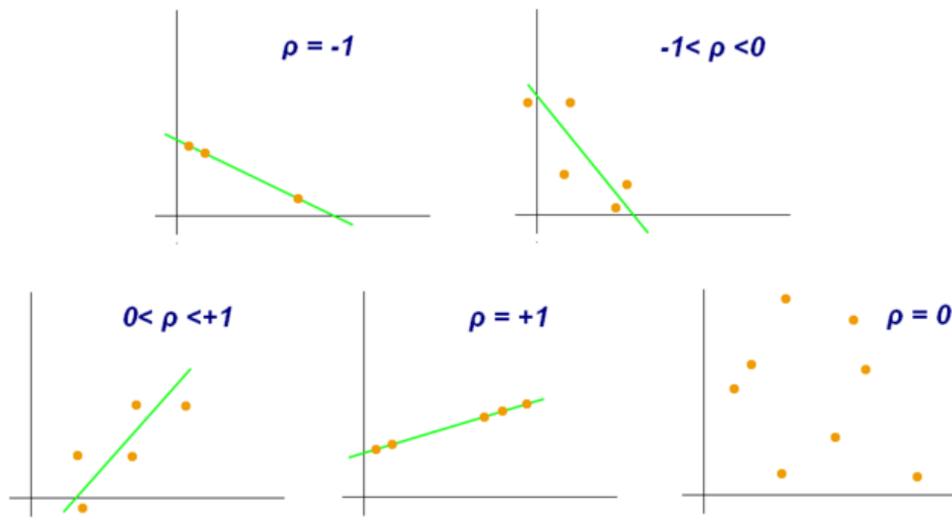


Figura 9: Cuando el coeficiente r de Pearson se eleva al cuadrado, el resultado indica el porcentaje de la variación de una variable debido a la variación de la otra y viceversa. Es decir, el coeficiente de determinación, r al cuadrado o  $r^2$ , es la proporción de la variación en Y explicada por X. Puede adoptar cualquier valor entre 0 y 1, inclusive.

**Implementación en Excel:** En Excel, esta fórmula se simplifica considerablemente mediante el uso de la función PEARSON. Por ejemplo, supongamos que tenemos el siguiente conjunto de datos en Excel que muestra la tasa de delincuencia y ciertos factores socioeconómicos relevantes para un análisis criminal:

	A	B
1	Mes	Tasa de Delincuencia
2	1	3.2
3	2	2.8
4	3	3.5
5	4	3.1
6	5	2.9
7	6	3.3

En Excel, colocaríamos la fórmula =PEARSON (B2 : B7, C2 : C7) en una celda para calcular la correlación de Pearson entre la tasa de delincuencia y los factores socioeconómicos relevantes. Supongamos que el resultado es  $r = 0,98$ , indicando una fuerte correlación positiva entre estas variables.

## 12. Regresión Lineal en Análisis Criminal

La regresión lineal es una técnica estadística que permite modelar la relación entre una variable dependiente y una o más variables independientes. En el contexto del análisis criminal, la regresión lineal puede ser utilizada para predecir, por ejemplo, futuras tasas de delincuencia basándose en variables como el desempleo, la educación, entre otras.

La fórmula general de la regresión lineal es:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

En Excel, la herramienta de Análisis de Regresión puede ser utilizada para realizar una regresión lineal. Esta herramienta se encuentra en el grupo *Análisis de Datos* de la pestaña *Datos*. Para utilizarla, es necesario seleccionar el rango de datos para las variables dependiente e independiente, y luego seguir las instrucciones del asistente para obtener los resultados de la regresión, incluyendo los coeficientes  $\beta_0$  y  $\beta_1$ , y otras estadísticas relevantes. Por ejemplo, suponiendo que los datos de la variable dependiente (tasas de delincuencia) están en la columna B, y los de la variable independiente (factores socioeconómicos) están en la columna C, usted debería seleccionar estos rangos en la herramienta de Análisis de Regresión para obtener el modelo de regresión lineal.

**Implementación en Excel:** Utilizando el mismo conjunto de datos, podríamos querer modelar la tasa de delincuencia en función de factores socioeconómicos utilizando regresión lineal. En Excel, seguiríamos estos pasos:

- Ir a la pestaña *Datos*.
- Seleccionar *Análisis de Datos* y luego *Regresión*.
- En el campo *Rango de entrada Y*, introducir el rango de la tasa de delincuencia (B2 : B7).
- En el campo *Rango de entrada X*, introducir el rango de los factores socioeconómicos (C2 : C7).
- Seguir las instrucciones y revisar los resultados proporcionados por Excel.

Excel nos proporcionará el siguiente gráfico:

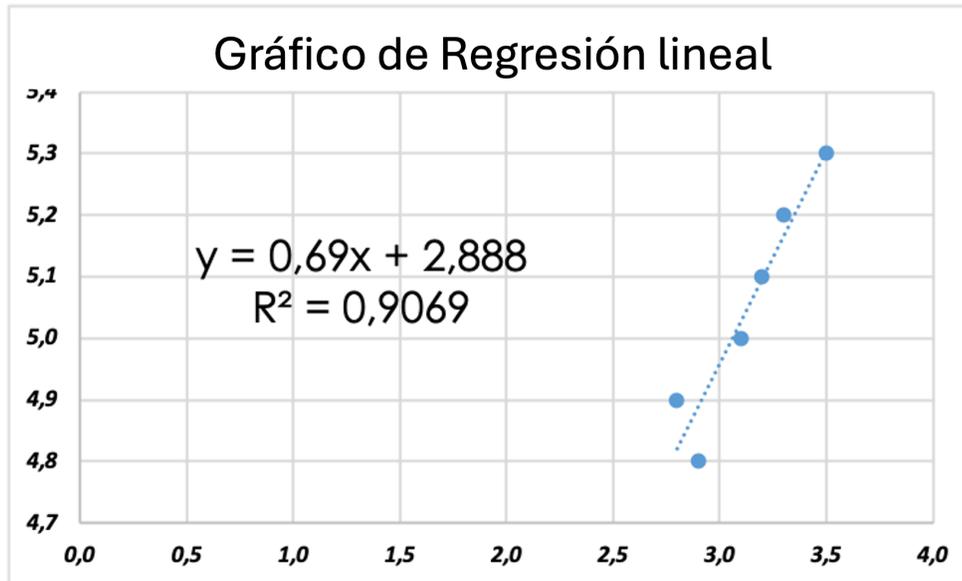


Figura 10: Cuando el coeficiente  $r$  de Pearson se eleva al cuadrado, el resultado indica el porcentaje de la variación de una variable debido a la variación de la otra y viceversa. Es decir, el coeficiente de determinación,  $r$  al cuadrado o  $r^2$ , es la proporción de la variación en la tasa de delincuencia explicada por el factor analizado. Puede adoptar cualquier valor entre 0 y 1, inclusive.

Luego reemplazamos los valores de las betas estimadas quedando los siguientes coeficientes para la regresión lineal:

$$\hat{\beta}_0 = 2,888, \quad \hat{\beta}_1 = 0,69$$

Entonces, la ecuación de la regresión lineal sería:

$$\hat{y} = 2,888 + 0,69x$$

Donde  $\hat{y}$  es la tasa de delincuencia estimada y  $x$  es el factor socioeconómico o demográfico.

“0.69 unidades, más una constante de 2.888” se refiere a los coeficientes obtenidos de la regresión lineal en el ejemplo proporcionado.

En una regresión lineal, la relación entre las variables se expresa con la ecuación  $y = \beta_0 + \beta_1x + \varepsilon$ , donde:

- $\beta_0$  es la constante o el intercepto,
- $\beta_1$  es el coeficiente de la variable independiente  $x$ ,
- $x$  es la variable independiente,
- $y$  es la variable dependiente (tasa de delincuencia),
- $\varepsilon$  es el error aleatorio.

En este caso,  $\beta_0 = 2,888$  y  $\beta_1 = 0,69$ . Esto implica que, según el modelo, para cada aumento de 1 unidad en el factor socioeconómico o demográfico, se espera que la tasa de delincuencia aumente en 0.69 unidades, además de la constante de 2.888.

Este resultado puede ser interpretado de la siguiente manera en la ecuación de la regresión lineal:

$$y = 2,888 + 0,69x + \varepsilon$$

En términos prácticos, si por ejemplo el valor del factor socioeconómico o demográfico es 10, entonces la tasa de delincuencia estimada, según este modelo, sería:

$$y = 2,888 + 0,69(10) = 2,888 + 6,9 = 9,788$$

Esto significa que con un valor de 10 para el factor socioeconómico o demográfico, se espera una tasa de delincuencia de 9.788 (sin tener en cuenta el error  $\varepsilon$ ).

El coeficiente de determinación,  $R^2 = 0,9069$ , sugiere que el 90,69 % de la variabilidad en la tasa de delincuencia puede ser explicada por este factor, según el modelo lineal propuesto. Esto sugiere que, en este ejemplo simplificado, hay una relación positiva fuerte entre el factor analizado y la tasa de delincuencia.

**Implementación de predicción:** Con el análisis de regresión lineal proporcionado, se tiene la siguiente ecuación de regresión lineal obtenida a partir de los datos:

$$y = 0,69x + 2,888$$

Donde:

- $y$  representa la Tasa de Delincuencia,
- $x$  representa un factor socioeconómico o demográfico relevante.

Para realizar una predicción dentro del rango de datos, se puede sustituir un valor específico de la variable independiente  $x$  en la ecuación de regresión lineal. Por ejemplo, para estimar la Tasa de Delincuencia cuando un factor socioeconómico o demográfico específico tiene un valor de 10, se realiza el siguiente cálculo:

$$y = 0,69(10) + 2,888 = 9,788$$

De acuerdo con esta estimación, cuando el factor socioeconómico o demográfico es 10, la Tasa de Delincuencia estimada es 9,788.

Este valor de 9,788 es una estimación de la Tasa de Delincuencia calculada a partir de la ecuación de regresión lineal, y está basada en un valor de 10 para el factor socioeconómico o demográfico. Según el modelo de regresión lineal proporcionado, cuando este factor es 10, la Tasa de Delincuencia estimada es 9,788.

Es importante tener en cuenta que este valor es una estimación modelada a partir de los datos con los que se construyó el modelo de regresión. La precisión de esta estimación depende de la calidad del

modelo y de la exactitud con que este modelo refleje la relación real entre el factor socioeconómico o demográfico y la Tasa de Delincuencia. Por lo tanto, siempre es bueno tener precaución al interpretar y utilizar las estimaciones obtenidas de un modelo de regresión lineal, especialmente si se está extrapolando fuera del rango de los datos originales.

El valor de  $R^2 = 0,9069$  indica que el modelo explica aproximadamente el 90,69 % de la variabilidad en la Tasa de Delincuencia, pero aún hay un 9,31 % de la variabilidad que no está explicada por el modelo. Por lo tanto, la interpretación y utilización de las estimaciones obtenidas del modelo deben ser realizadas con cautela, teniendo en cuenta la posibilidad de otros factores no considerados en el modelo que también pueden afectar la Tasa de Delincuencia. Estos análisis y modelos pueden proporcionar información valiosa a los analistas delictuales para entender mejor los factores que contribuyen a la delincuencia y desarrollar estrategias efectivas para su prevención y control.

### 13. Proceso de identificación de Patrones Delictuales en Excel

Identificar patrones delictuales es un aspecto crucial en el análisis delictual, ya que permite a los profesionales y a sus jefaturas anticipar, entender y, en última instancia, abordar los problemas de delincuencia de una manera más informada. A continuación se proporciona un desglose detallado de cómo se pueden identificar patrones delictuales utilizando Excel:



Figura 11: El análisis de datos implica examinar sets de datos para extraer conclusiones que informen decisiones o expandan el conocimiento sobre temas variados. Mediante diversas operaciones, no predefinidas debido a posibles desafíos en la recolección de datos, se busca obtener insights precisos para alcanzar objetivos establecidos.

#### 13.1. Recolección y Preparación de Datos:

- **Importación de Datos:** Importe los datos delictuales en Excel. Los datos pueden incluir detalles como la fecha, hora, ubicación, tipo de delito, entre otros.

- **Limpieza de Datos:** Asegúrese de que los datos estén limpios, consistentes y sin valores faltantes o errores, utilizando funciones como `ESPACIOS`, `LIMPIAR`, o la opción `ELIMINAR DUPLICADOS` desde la pestaña de datos.

### 13.2. Análisis Exploratorio de Datos (EDA):

- **Tablas Dinámicas:** Utilice tablas dinámicas para obtener un resumen de los datos y identificar tendencias iniciales, como la frecuencia de ciertos tipos de delitos o la variación de delitos en diferentes áreas o períodos de tiempo.
- **Funciones Estadísticas:** Utilice funciones estadísticas básicas como `CONTAR.SI`, `SUMAR.SI`, `PROMEDIO.SI` para obtener estadísticas descriptivas que puedan revelar patrones.

### 13.3. Visualización de Datos:

- **Gráficos de Tendencias:** Cree gráficos que muestren cómo varían los delitos a lo largo del tiempo (diariamente, mensualmente, anualmente).
- **Mapas de Calor:** Si tiene datos geográficos, los mapas de calor pueden ser útiles para identificar áreas de alta criminalidad.
- **Gráficos de Barras/Columnas:** Para comparar la frecuencia de diferentes tipos de delitos o delitos en diferentes áreas.

### 13.4. Análisis Temporal y Espacial:

- **Análisis Temporal:** Examine cómo los delitos varían a lo largo del tiempo, buscando patrones diarios, semanales, mensuales o estacionales.
- **Análisis Espacial:** Analice la distribución geográfica de los delitos para identificar zonas de alta criminalidad o patrones de desplazamiento.

### 13.5. Búsqueda y Análisis de Patrones Específicos:

- **Funciones de Búsqueda:** Utilice funciones como `BUSCARV`, `BUSCARH`, `INDICE` y `COINCIDIR` para buscar y analizar datos específicos.
- **Análisis de Cluster:** Aunque Excel no es la mejor herramienta para esto, con algunos complementos o software adicional, puede realizar análisis de clúster para identificar grupos de delitos similares.

### 13.6. Correlación y Regresión:

- Identifique si hay alguna correlación entre diferentes variables (ej: clima, eventos especiales, desempleo) y la incidencia delictiva utilizando la función `COEF.DE.CORREL` para la correlación y la herramienta de Análisis de Regresión para el análisis de regresión lineal.

### 13.7. Interpretación y Comunicación de Resultados:

Una vez identificados los patrones, es crucial interpretar los resultados en un contexto delictual y comunicar los hallazgos a las partes interesadas para tomar decisiones informadas. La comunicación efectiva de los resultados analíticos es un arte en sí mismo y juega un papel fundamental en cómo se perciben y utilizan los hallazgos. Una manera excepcional de mejorar la comunicación y visualización de los resultados es mediante el uso de paneles de control.

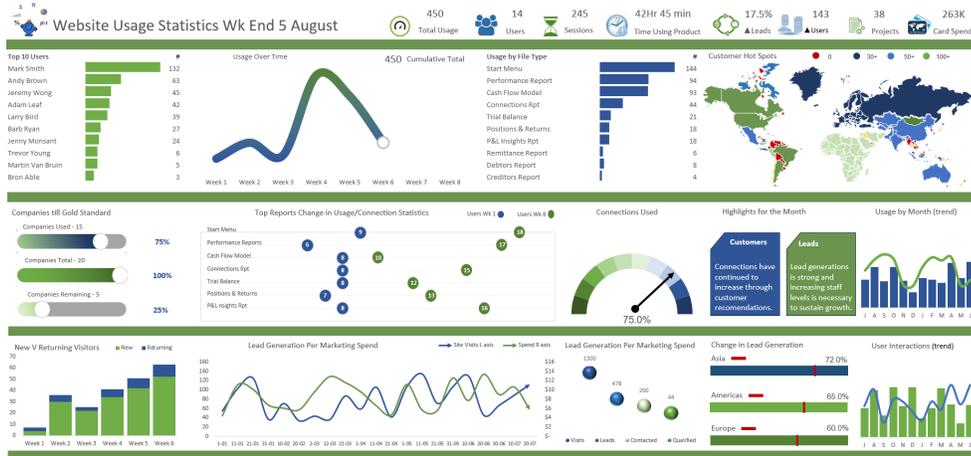


Figura 12: Un tablero bien diseñado puede proporcionar una vista clara y comprensible del análisis delictual, facilitando a las partes interesadas la digestión de la información y la toma de decisiones basadas en datos.

En el ámbito de la analítica criminal, un tablero puede ser configurado para mostrar varias métricas clave, tendencias y patrones geospaciales, distribuciones temporales de incidentes delictuales y cualquier clasificación o análisis de regresión que se haya realizado. Además, los paneles interactivos permiten a los usuarios explorar los datos por sí mismos, lo que puede ser invaluable para fomentar una comprensión más profunda y una discusión enriquecedora sobre los hallazgos.

Por ejemplo, un tablero puede mostrar un mapa de calor de la actividad delictiva junto con gráficos de tendencias temporales, permitiendo a las partes interesadas visualizar cómo la delincuencia se distribuye tanto espacial como temporalmente. Los filtros interactivos pueden permitir a los usuarios ajustar el período de tiempo o el tipo de delito mostrado, proporcionando así una herramienta poderosa para explorar cómo diferentes factores pueden influir en los patrones delictuales.

La capacidad de un tablero para consolidar múltiples fuentes de datos y presentar análisis complejos de manera visual y accesible hace que sea una herramienta indispensable para los analistas criminales. Al adoptar paneles de control como una herramienta de comunicación, los analistas criminales pueden mejorar significativamente la forma en que presentan y explican sus hallazgos, facilitando una discusión informada y basada en datos con las partes interesadas. Esto, a su vez, contribuye a una toma de decisiones más eficaz y una estrategia de prevención y respuesta al delito más informada y efectiva, por parte de los gestores ya sea de Carabineros de Chile, PDI o Encargados de Seguridad Municipales.

### 13.8. Recomendaciones:

- El dominio de Excel representa el punto de partida esencial en el arsenal de habilidades de un analista criminal profesional. Esta herramienta proporciona una base sólida en el manejo y análisis básico de datos. Sin embargo, para ascender al siguiente nivel de profesionalismo y abordar análisis más complejos y detallados, es crucial orientarse hacia el uso de software especializado en análisis estadístico y geoespacial, como Python, R o GIS.

Este proceso, que comienza con el uso experto de Excel, sienta las bases para que los analistas criminales identifiquen patrones delictuales de manera eficaz. Al dominar Excel, los analistas pueden recopilar, organizar y analizar datos de manera que se revelen tendencias y correlaciones clave, proporcionando así información valiosa para la toma de decisiones. Este nivel de detalle y precisión es fundamental para apoyar las estrategias de prevención y gestión de la delincuencia, permitiendo una respuesta más informada y dirigida.

A medida que los analistas criminales progresan en su carrera, la transición hacia herramientas más avanzadas les permite ampliar y profundizar su análisis. Este salto cualitativo en la capacidad analítica abre nuevas puertas para la interpretación de datos, permitiendo no solo una comprensión más completa de los patrones delictuales, sino también facilitando el diseño de estrategias más efectivas y personalizadas para combatir el crimen. La integración de análisis estadístico avanzado, modelado predictivo y mapeo geoespacial, por ejemplo, posibilita una visión más holística y multidimensional del panorama delictivo, lo que resulta esencial en el contexto actual de complejidad creciente de los patrones delictuales.

## Referencias

- [1] Steven Johnson. *The Ghost Map: The Story of London's Most Terrifying Epidemic—and How It Changed Science, Cities, and the Modern World*. Riverhead Books, 2006.
- [2] John Snow. *On the Mode of Communication of Cholera*. Wilson and Ogilvy, 1855.
- [3] E. W. Gilbert. Pioneer maps of health and disease in england. *The Geographical Journal*, 124(2):172–183, 1958.
- [4] Tom Koch and Kenneth Denike. *Spatial Analysis in Epidemiology: with a focus on Geographical Information Systems*. Oxford University Press, 2011.
- [5] Jordan Ellenberg. *How Not to Be Wrong: The Power of Mathematical Thinking*. Penguin Books, 2014.
- [6] James Q. Wilson and George L. Kelling. *Broken Windows: The police and neighborhood safety*. 1982.
- [7] Paul J. Brantingham and Patricia L. Brantingham. *Environmental Criminology*. Waveland Press, 1981.
- [8] Lawrence W. Sherman, Patrick R. Gartin, and Michael E. Buerger. Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology*, 27(1):27–56, 1989.
- [9] Nassim Nicholas Taleb. *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.

## **Sobre el Autor:**

Con un Magíster en Data Science de la Universidad San Sebastián y avanzando hacia un Doctorado en la misma disciplina en la Universidad Adolfo Ibáñez, el autor fusiona una robusta experiencia en Machine Learning, análisis estadístico y modelado predictivo con una sólida base en gestión empresarial adquirida como Ingeniero en Administración de Empresas y Control de Gestión. Su enfoque interdisciplinario ha impulsado proyectos transversales que mejoran la toma de decisiones, abarcando desde la seguridad hasta la optimización de recursos en varios sectores.

Con una distinguida trayectoria de 25 años en Carabineros, se ha especializado en áreas críticas como derechos humanos, sistemas de información, gestión estratégica, prevención del delito basada en análisis criminal y estadística aplicada al análisis criminal. Su trabajo académico culmina en la publicación 'Anticipación de los niveles de prioridad policial en grupos delictivos utilizando Classifier Ensemble', que propone un modelo innovador para predecir áreas críticas de vigilancia, contribuyendo significativamente a la eficiencia de las operaciones de seguridad tanto en el ámbito público como privado.



## **Resumen:**

Este dossier está diseñado para complementar y enriquecer las habilidades de los analistas criminales en Carabineros de Chile, la Policía de Investigaciones (PDI) y en áreas de seguridad municipales. Su propósito es apoyar y mejorar las prácticas existentes, sin competir con los procesos de formación y aplicabilidad de cada institución. Ofrecemos este material como un recurso adicional, para integrarse efectivamente en los métodos actuales y fortalecer la lucha continua contra el crimen, respetando siempre los protocolos y directrices de cada entidad.