



**UNIVERSIDAD NACIONAL DE CHIMBORAZO**

**VICERRECTORADO DE INVESTIGACIÓN,**

**VINCULACIÓN Y POSGRADO**

**DIRECCIÓN DE POSGRADO**

**TESIS PREVIA A LA OBTENCIÓN DEL GRADO DE:**

**MAGÍSTER EN MATEMÁTICA APLICADA MENCIÓN MATEMÁTICA  
COMPUTACIONAL**

**TEMA:**

**“DETECCIÓN DE PATRONES DE TIEMPO Y ESPACIO EN EL SERVICIO  
INTEGRADO DE SEGURIDAD ECU 911 PARA IDENTIFICAR LAS PARROQUIAS  
DE ALTA PELIGROSIDAD EN LA CIUDAD DE QUEVEDO.”**

**AUTOR:**

**Ing. WILMER FABRICIO ALMEIDA MURILLO**

**TUTOR:**

**Ing. HENRY MAURICIO VILLA YÁNEZ, Mgs**

**Riobamba, Ecuador. 2025**

## Certificación del Tutor

Certifico que el presente trabajo de titulación denominado: **“Detección de patrones de tiempo y espacio en el servicio integrado de seguridad ECU 911 para identificar las parroquias de alta peligrosidad en la ciudad de Quevedo”** ha sido elaborado por el Ingeniero Wilmer Fabricio Almeida Murillo, el mismo que ha sido orientado y revisado con el asesoramiento permanente de mi persona en calidad de Tutor. Así mismo, refrendo que dicho trabajo de titulación ha sido revisado por la herramienta antiplagio institucional; por lo que certifico que se encuentra apto para su presentación y defensa respectiva.

Es todo cuanto puedo informar en honor a la verdad.

Riobamba, 24 de febrero de 2025



Firmado electrónicamente por:  
HENRY MAURICIO  
VILLA YANEZ

---

Ing. Henry Mauricio Villa Yáñez, Mgs

**TUTOR**

## Declaración de Autoría y Cesión de Derechos

Yo, **Wilmer Fabricio Almeida Murillo** con número único de identificación **092897688-5**, declaro y acepto ser responsable de las ideas, doctrinas, resultados y lineamientos alternativos realizados en el presente trabajo de titulación denominado: “Título del trabajo de titulación.” previo a la obtención del grado de **Magíster en Matemática Aplicada con mención en Matemática Computacional**.

- Declaro que mi trabajo investigativo pertenece al patrimonio de la Universidad Nacional de Chimborazo de conformidad con lo establecido en el artículo 20 literal j de la Ley Orgánica de Educación Superior LOES.
- Autorizo a la Universidad Nacional de Chimborazo que pueda hacer uso del referido trabajo de titulación y a difundirlo como estime conveniente por cualquier medio conocido, y para que sea integrado en formato digital al Sistema de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor, dando cumplimiento de esta manera a lo estipulado en el artículo 144 de la Ley Orgánica de Educación Superior LOES.

Riobamba, 24 de febrero de 2025



Firmado electrónicamente por:  
**WILMER FABRICIO  
ALMEIDA MURILLO**

---

Ing. Wilmer Fabricio Almeida Murillo

N.U.I. 092897688-5



Dirección de  
Posgrado  
VICERRECTORADO DE INVESTIGACIÓN,  
VINCULACIÓN Y POSGRADO



Riobamba, 18 de febrero de 2025

## ACTA DE SUPERACIÓN DE OBSERVACIONES

En calidad de miembro del Tribunal designado por la Comisión de Posgrado, CERTIFICO que una vez revisado el Proyecto de Investigación y/o desarrollo denominado "**DETECCIÓN DE PATRONES DE TIEMPO Y ESPACIO EN EL SERVICIO INTEGRADO DE SEGURIDAD ECU 911 PARA IDENTIFICAR LAS PARROQUIAS DE ALTA PELIGROSIDAD EN LA CIUDAD DE QUEVEDO.**", dentro de la línea de investigación de **Modelos matemáticos, físicos, continuos, discretos, determinísticos, estocásticos**, presentado por el maestrante **ALMEIDA MURILLO WILMER FABRICIO**, portador de la Ci. 0928976885, del programa de **MAESTRÍA EN MATEMÁTICA APLICADA MENCIÓN MATEMÁTICA COMPUTACIONAL**, cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo lo que podemos certificar en honor a la verdad.

Atentamente,



Firmado electrónicamente por:  
**HENRY MAURICIO  
VILLA YÁNEZ**

Mgs. Henry Mauricio Villa  
Yáñez  
**TUTOR**



Firmado electrónicamente por:  
**GUILLERMO EDVIN  
MACHADO SOTOMAYOR**

Phd. Guillermo Edvin  
Machado Sotomayor Mgs.  
**MIEMBRO DEL TRIBUNAL**



Firmado electrónicamente por:  
**PAOLA GABRIELA  
VINUEZA NARANJO**

Phd. Paola Gabirela  
Vinueza Naranjo Mgs.  
**MIEMBRO DEL TRIBUNAL**



Campus La Dolorosa  
Av. Eloy Alfaro y 10 de Agosto  
Teléfono (593-3) 373-0880, ext. 2002  
Riobamba - Ecuador

**Unach.edu.ec**  
*en movimiento*

## CERTIFICADO

De mi consideración:

Yo, **Henry Mauricio Villa Yáñez**, certifico que el estudiante Almeida Murillo Wilmer Fabricio, con cédula de identidad No. **0928976885**, ha presentado su trabajo de titulación dentro del programa de **Maestría en Matemática Aplicada con mención en Matemática Computacional**, correspondiente a la Tercera Cohorte (2024-2025). El trabajo fue desarrollado bajo la modalidad de Proyecto de Titulación con componente de investigación aplicada/desarrollo, denominado: **“Detección de patrones de tiempo y espacio en el servicio integrado de seguridad ECU 911 para identificar las parroquias de alta peligrosidad en la ciudad de Quevedo.”**, el mismo fue sometido al sistema de verificación de similitud de contenido **Compilatio**, obteniendo un **1%** de coincidencia en el texto.

Por lo expuesto, extiendo el presente certificado en honor a la verdad.

Riobamba, 24 de febrero de 2025



Firmado electrónicamente por:  
**HENRY MAURICIO  
VILLA YANEZ**

---

Ing. Henry Mauricio Villa Yáñez, Mgs

TUTOR

Adj.-

- Resultado de análisis de similitud

## **Agradecimiento**

Agradezco profundamente a Dios por darme la fuerza y sabiduría para llegar hasta aquí. A mis padres, por su amor, apoyo y sacrificio incondicional, que siempre han sido una fuente de inspiración. Mi sincero agradecimiento a mis docentes de la maestría por su valiosa orientación a lo largo de este proceso. A mi tutor, el Mgs. Henry Villa, por su constante ayuda y dedicación en la elaboración de este documento. También a mi esposa, por su amor, paciencia y apoyo constante, y a mis hermanos, por su motivación y respaldo. Sin su colaboración, este trabajo no habría sido posible.

Wilmer Almeida

## **Dedicatoria**

Dedico este trabajo a mis hijos, quienes son mi mayor fuente de inspiración y motivación. Su amor y alegría me impulsan a seguir adelante, siempre buscando un futuro mejor para ellos. A mi esposa, por su apoyo incondicional, amor y paciencia, que me han acompañado en cada paso de este proceso. Gracias por ser mi compañera en todos los aspectos de la vida. También, dedico este esfuerzo a la ciencia, que, con su constante búsqueda del conocimiento y la verdad, transforma el mundo y nos permite crecer como seres humanos. Que este trabajo sea una pequeña contribución al vasto campo del conocimiento, con la esperanza de que inspire a otros a seguir explorando y aprendiendo.

## Índice General

<b>Certificación del Tutor .....</b>	<b>ii</b>
<b>Declaración de Autoría y Cesión de Derechos .....</b>	<b>iii</b>
<b>Acta de Superación de Observaciones .....</b>	<b>iv</b>
<b>Certificado .....</b>	<b>v</b>
<b>Agradecimiento .....</b>	<b>vi</b>
<b>Dedicatoria .....</b>	<b>vii</b>
<b>Índice General.....</b>	<b>viii</b>
<b>Índice de Tablas .....</b>	<b>xi</b>
<b>Índice de Figuras .....</b>	<b>xii</b>
<b>Resumen .....</b>	<b>xiii</b>
<b>Abstract .....</b>	<b>xiv</b>
<b>Introducción.....</b>	<b>15</b>
<b>Capítulo 1 Generalidades.....</b>	<b>17</b>
1.1. Planteamiento del Problema .....	17
1.2. Justificación .....	18
1.3. Objetivos.....	20
1.3.1. Objetivos General.....	20
1.3.2. Objetivos Específicos .....	20
<b>Capítulo 2 Estado del Arte y la Práctica .....</b>	<b>22</b>
2.1. Antecedentes Investigativos .....	22
2.2. Fundamentación Legal.....	26
2.3. Fundamentación Teórica .....	27
2.3.1. Seguridad Ciudadana .....	27
2.3.1.1. Sistema Integrado de Seguridad ECU 911 .....	27
2.3.1.2. Problemas Actuales de Seguridad en Quevedo .....	28
2.3.1.3. Estrategia de Gestión del ECU 911 .....	29
2.3.1.4. Coordinación Interinstitucional en la Gestión de Emergencias.....	29



2.3.2. Patrones espaciales y temporales .....	30
2.3.2.1. Clasificación de patrones espaciales.....	31
2.3.2.2. Métodos para identificar patrones delictivos.....	32
2.3.2.3. Factores que influyen en los Patrones Delictivos .....	34
2.3.3. Herramientas y Métodos para el Análisis de Patrones .....	34
2.3.3.1. Sistema de Información Geográfico (SIG) en la Seguridad Pública .....	34
2.3.3.2. Mapas de Calor .....	35
2.3.3.3. Técnicas de Minería de Datos para la Seguridad .....	36
<b>Capítulo 3 Diseño Metodológico.....</b>	<b>41</b>
3.1. Enfoque de la Investigación.....	41
3.2. Diseño de la Investigación.....	41
3.3. Tipo de Investigación .....	41
3.4. Nivel de Investigación .....	42
3.5. Técnicas e Instrumentos de Recolección de Datos.....	42
3.6. Técnicas para el Procesamiento e Interpretación de Datos .....	42
3.7. Población y Muestra .....	43
3.7.1. Población.....	43
3.7.2. Tamaño de la Muestra.....	43
<b>Capítulo 4 Análisis Y Discusión de los Resultados .....</b>	<b>46</b>
4.1. Análisis Descriptivo de los Resultados.....	46
4.1.1. Organización de los Datos .....	46
4.1.1.1. Incidentes por Circuito .....	47
4.1.1.2. Incidentes por Subcircuito .....	49
4.1.1.3. Horas Pico con Mayor Cantidad de Incidentes .....	50
4.1.1.4. Incidentes más Frecuentes .....	51
4.1.1.5. Relación entre Circuitos e Incidentes .....	52
4.1.2. Herramientas y Técnicas Utilizadas.....	53
4.1.3. Análisis Espacial.....	55
4.1.3.1. Aplicación de Algoritmos de Clústering. ....	58
4.1.3.2. Evaluación de Algoritmos. ....	63
4.1.4. Análisis Temporal.....	66
4.1.4.1. Mapa de Calor de los Incidentes por Día de la Semana y Hora del Día. ....	66

4.1.5. Incidentes en Circuitos en Horas Pico .....	67
4.1.6. Evolución Mensual .....	68
4.1.7. Mapa de Calor.....	70
4.2. Discusión de los Resultados .....	72
<b>Capítulo 5 Marco Propositivo .....</b>	<b>75</b>
5.1. Planificación de la Actividad Preventiva.....	75
Conclusiones.....	77
Referencias Bibliográficas.....	79
Apéndice.....	82
Apéndice A. líneas de código usadas para el análisis de los datos.....	82

## Índice de Tablas

Tabla 1. Factores que influyen en los patrones delictivo .....	34
Tabla 2. Aspectos del Sistemas de Información Geográfica .....	35
Tabla 3. Código de los circuitos y subcircuitos del cantón Quevedo.....	47
Tabla 4. Incidente por circuito.....	48
Tabla 5. Incidente por subcircuito .....	49
Tabla 6. Horas pico.....	50
Tabla 7. Incidentes más frecuentes.....	51
Tabla 8. Resultados obtenidos de la métrica .....	64
Tabla 9. Comparativo e interpretacion de los resultados .....	65

## Índice de Figuras

Figura 1. Modelo de gestión sistema integrado de seguridad ECU 911 .....	28
Figura 2. Estrategia de gestión .....	29
Figura 3. Coordinación de Interinstitucional de emergencias .....	30
Figura 4. Mapa de calor .....	36
Figura 5. Arbol de decisión .....	40
Figura 6. Horas pico con mayores incidentes.....	50
Figura 7. Top 15 de incidentes .....	52
Figura 8. Mapa de calor de distribución de incidentes por circuito y tipo de incidente en Quevedo .....	53
Figura 9. Incidentes por parroquia.....	55
Figura 10. Distribución geográfica de los incidentes. ....	56
Figura 11. Distribución de puntos en la ciudad .....	57
Figura 12. Gráfica algoritmo Kmeans .....	58
Figura 13. Centroide Kmeans.....	59
Figura 14. Algoritmo DBSCAN.....	60
Figura 15. Algoritmo HDBSCAN.....	61
Figura 16. HDBSCAN clústers significativos.....	62
Figura 17. Incidentes por dia de la semana y hora del día.....	67
Figura 18. Incidentes en circuitos en horas pico .....	68
Figura 19. Evolución Mensual de los incidentes.....	69
Figura 20. Mapa de calor puntos calientes .....	70

## Resumen

Este trabajo de investigación se enfoca en la detección de patrones temporales y espaciales en el servicio integrado de seguridad ECU 911 Quevedo mediante técnicas de asociación de datos, con el objetivo de identificar parroquias de alta peligrosidad. Para ello, se utiliza información proporcionada por el servicio de seguridad de la zona Babahoyo, que registra todas las incidencias ocurridas en la provincia de Los Ríos durante el año 2023, año en el que se registraron un total de 3,447,071 llamadas a nivel nacional. Representando un promedio de manera diaria de 9,444 avisos (lo que corresponde a 393 eventos que se suscitan por hora) y una frecuencia mensual de 287,256 incidentes reportados. Teniendo que, el 70% corresponde a seguridad ciudadana, el 12% a gestión sanitaria, y el 11% a tránsito y movilidad, por último, el 7% restante se distribuye entre categorías como servicios municipales, gestión de siniestros, servicio militar y gestión de riesgos

Mediante el uso de algoritmos de clustering, los datos generados en cada interacción se agrupan en conjuntos de eventos similares, denominados clústeres, lo que permite la identificación de patrones ocultos. Este enfoque proporciona una visibilidad detallada de las situaciones que ocurren en zonas específicas, facilitando la toma de decisiones por parte de las entidades colaboradoras en seguridad ciudadana. Estas decisiones pueden contribuir a la reducción de costos y tiempos de respuesta ante emergencias.

**Palabras claves:** Patrones de tiempo, análisis de datos, algoritmos, clustering, seguridad

**Abstract**

This research focuses on detecting temporal and spatial patterns in the integrated security service ECU 911 Quevedo through data association techniques to identify highly dangerous parishes. This information provided by the security service of the Babahoyo area is used, which records all incidents that occurred in the province of Los Ríos during the year 2023, the year in which a total of 3,447,071 calls were recorded nationwide. Representing a daily average of 9,444 notices (393 events per hour) and a monthly frequency of 287,256 reported incidents. Considering that 70% corresponds to citizen security, 12% to health management, and 11% to traffic and mobility, the remaining 7% is distributed between municipal services, accident management, military service, and risk management categories.

With clustering algorithms, the data generated in each interaction is grouped into similar events called clusters to identify hidden patterns. This approach provides detailed visibility of situations in specific areas, facilitating decision-making by collaborating entities in citizen security. These decisions can contribute to reducing costs and response times in emergencies.

**Keywords:** Temporal patterns, data analysis, algorithms, clustering, security.



ANA ELIZABETH  
MALDONADO LEÓN

Reviewed by:

Ms.C. Ana Maldonado León

ENGLISH PROFESSOR

C.I.0601975980

## 1. Introducción

La detección de patrones espacio-temporales es una técnica esencial a utilizar en la toma de decisiones en materia de seguridad, debido a que permite identificar zonas y momentos críticos donde se suscitan eventos de riesgo. En el contexto de la ciudad de Quevedo, el análisis de los datos anuales proporcionados por la entidad pública del servicio integrado de seguridad ECU 911 se convierte en una herramienta clave para identificar las parroquias con mayores niveles de peligrosidad, lo que permite diseñar estrategias preventivas y operativas más eficientes. (Ortiz, L., & Ponce, M. , 2020)

El presente estudio tiene como objetivo la detección de patrones espacio-temporales en las llamadas y reportes de emergencias registradas por el ECU 911 Quevedo durante el periodo enero a diciembre del año 2023. Mediante el uso de técnicas de análisis de datos espaciales y series temporales, se busca identificar los intervalos de tiempo durante el día y las áreas geográficas donde se concentra un mayor número de incidentes. Estos hallazgos permitirán no solo una mejor distribución de los recursos de seguridad, sino también la elaboración de políticas preventivas que permitan una mayor efectividad en la ejecución de acciones de emergencia y contribuya a reducir los riesgos en las zonas más vulnerables. (Torres, F., & Gutiérrez, P., 2021)

Para ello, se emplearán metodologías de análisis geoespacial como mapas de calor, y modelos de regresión espacio-temporal, complementados con herramientas de series temporales para estudiar la evolución temporal de los incidentes. (Díaz, C., & Hernández, R., 2019)

El objetivo es que esta investigación es contribuir al diseño de políticas más efectivas que ayuden a gestionar emergencias de manera eficiente, brindando mayor seguridad y mejorando la calidad de vida de la comunidad. Además, el uso de tecnologías como la inteligencia artificial y el análisis de datos en los servicios públicos no solo moderniza la

seguridad ciudadana, sino que también permite una respuesta más rápida y acertada ante situaciones de riesgo

Debido a la capacidad de agrupar los datos se utilizan algoritmos de clústeres, como Kmeans, optimizando la identificación de áreas con mayores incidencias de emergencias, por otra parte, DBSCAN y HDBSCAN ayudan en la identificación de clústeres de manera flexible

En cuanto a los algoritmos específicos, K-Means se utilizará por su capacidad para agrupar datos en clústeres definidos, optimizando la identificación de áreas de alta incidencia. DBSCAN y HDBSCAN, por su parte, permiten identificar clústeres de forma más flexible, se adaptan mejor a los cambios en los datos, incluso cuando aparecen valores fuera de lo común. Gracias a estas técnicas, es posible manejar grandes cantidades de información y detectar patrones delictivos con mayor precisión y rapidez.

El análisis se enfoca en las parroquias tanto urbanas como rurales de la ciudad de Quevedo, con especial atención a la distribución y concentración de incidentes de alto riesgo, como lo son los robos, accidentes y disturbios. Según Gómez y Ramírez mantienen en su investigación que el análisis conjunto de patrones temporales y espaciales es clave para implementar mejoras en las estrategias de respuesta ante los incidentes en áreas metropolitanas, lo que resulta de gran importancia para nuestra investigación en la ciudad de Quevedo. (Gómez, M., & Ramírez, S., 2023)

Este análisis no solo permitirá determinar los puntos críticos en cuanto a peligrosidad, sino que también facilitará la toma de decisiones basadas en datos para la optimización de los recursos de seguridad pública en Quevedo.



# Capítulo 1

## Generalidades

### 1.1. Planteamiento del Problema

En la ciudad de Quevedo, se observa un considerable incremento de eventos de inseguridad, tales como, robos, accidentes de tránsito, libadores y actos de violencia. El sistema de seguridad integrado ECU 911, es el encargado de la gestión de emergencias, reuniendo datos valiosos sobre estos incidentes, incluyendo el tiempo, tipología y lugar en el que ocurren. No obstante, esta información no ha sido utilizada de manera integral para identificar patrones de riesgo que permitan orientar de forma eficaz la distribución de los recursos en el aspecto de seguridad hacia la ciudadanía Quevedeña. Una adecuada interpretación de los patrones espacio-temporales puede mejorar considerablemente la asignación de recursos en áreas consideradas como vulnerables y críticas. (Torres, F., & Gutiérrez, P., 2021)

El principal desafío radica en que, ante la falta de una identificación adecuada de los patrones espacio-temporales de los incidentes, los recursos destinados a la seguridad y la atención de emergencias se asignan de manera reactiva y poco eficiente, lo que resulta en la desatención de áreas críticas en momentos clave. Como señalan (Díaz, C., & Hernández, R., 2019) la falta de modelos predictivos impide que las autoridades anticipen de manera precisa los incidentes en determinadas áreas geográficas. En consecuencia, las autoridades locales enfrentan dificultades para priorizar intervenciones en las parroquias con mayor nivel de peligrosidad y no pueden predecir con precisión cuándo y dónde ocurrirán más emergencias.

El reto consiste en que, a pesar de la gran cantidad de datos disponibles sobre los incidentes reportados por el ECU 911, no se han aplicado metodologías avanzadas que analicen los datos disponibles para identificar patrones específicos en el desarrollo de los eventos de emergencia. Es fundamental determinar con precisión cuáles son las parroquias

que tienen mayor riesgo y en qué horarios estas presentan mayor frecuencia de emergencias, utilizando herramientas de análisis geoespacial y temporal. (Gómez, M., & Ramírez, S., 2023)

Este estudio plantea la siguiente pregunta de investigación: ¿Cómo pueden los algoritmos de clustering ayudar a identificar patrones ocultos en los datos de incidentes del ECU 911 para detectar parroquias de alta peligrosidad en Quevedo y mejorar la eficiencia en la gestión de emergencias?

El planteamiento del problema resalta la necesidad de aplicar enfoques basados en datos para la prevención y atención de emergencias en Quevedo, lo que permitirá optimizar la seguridad y reducir el riesgo en las parroquias más vulnerables (López, A., & Vega, D., 2022)

## **1.2. Justificación**

La seguridad ciudadana es uno de los pilares fundamentales para el bienestar de cualquier sociedad, especialmente al referirnos a zonas urbanas de alta densidad poblacional como el de la ciudad de Quevedo. El aumento de los incidentes delincuenciales y de emergencias que se suscitan en la ciudad resalta la urgencia de mejorar la eficiencia en la asignación de recursos para las funciones de seguridad. Aunque, la entidad pública ECU 911 recopile datos relevantes sobre estos eventos, la ausencia de un análisis profundo de la información disponible, y de los patrones espacio-temporales genera que se limite la capacidad de las autoridades para optimizar la gestión y los tiempos de respuesta ante emergencias. El análisis de los patrones permite identificar las áreas de mayor riesgo y mejorar la planificación estratégica de los recursos disponibles. (López, A., & Ramírez, J., 2020)

Sin embargo, la entidad ECU 911 recoge datos importantes que se generan día a día sobre estos eventos, sin realizar un análisis de manera detallada para determinar los patrones

en el tiempo y el espacio dados en efecto durante el desarrollo de los incidentes, lo que dificulta que las autoridades puedan gestionar y responder a las emergencias de forma más eficiente.

La adopción de herramientas para el análisis geoespacial y temporal no solo contribuirá a identificar las parroquias con mayores riesgos, sino que también permitirá anticipar los períodos más propensos a la ocurrencia de emergencias. Esto permitirá asignar los recursos de manera más eficiente desde el inicio, mejorando la capacidad de respuesta ante situaciones críticas. Según lo señalado por Sánchez y Moreno (2021), los modelos predictivos tienen el potencial de reducir considerablemente los tiempos de respuesta y aumentar la eficiencia en los operativos de seguridad pública.

Implementar estas tecnologías en las 9 parroquias de la ciudad de Quevedo contribuirá tanto a mejorar la seguridad de los ciudadanos como a optimizar el uso de recursos disponibles, como las patrullas, ambulancias, vehículos de bomberos y el personal de emergencia necesario en la atención de cada evento. (Sánchez, P., & Moreno, F., 2021)

Además, la anticipación de incidentes mediante la detección de patrones es una de las formas más efectivas de prevenir la criminalidad y mejorar la seguridad en áreas urbanas, (Torres, L., & Gutiérrez, S., 2022). La capacidad de actuar de manera proactiva, en lugar de reactiva, permite a las autoridades reducir la incidencia de delitos en las zonas más afectadas y mejorar la percepción de seguridad entre los ciudadanos, lo cual es clave para fortalecer la confianza en las instituciones públicas (Pérez, D., & Delgado, C., 2019).

En conclusión, este estudio se justifica plenamente debido a la urgente necesidad de mejorar la seguridad en el territorio urbano y rural de la ciudad de Quevedo utilizando técnicas avanzadas de análisis de datos tanto temporales como espaciales. Al implementar algoritmos de clustering y tener un enfoque basado en los datos, no solo se podrá identificar más fácilmente las parroquias con mayor riesgo, sino también distribuir los recursos de

manera más eficiente, lo que podría reducir significativamente los delitos y mejorar la calidad de vida de los habitantes

Desde una perspectiva social, la adopción de estos métodos contribuiría a crear un entorno más seguro, disminuyendo la percepción de inseguridad entre los ciudadanos. Esto podría promover un mayor bienestar y cohesión social, al brindar a la comunidad la confianza de que las autoridades están gestionando la situación de manera eficiente y proactiva.

Económicamente, el uso del análisis espacio-temporal podría generar ahorros significativos en la gestión de emergencias. Al identificar los patrones de riesgo y predecir los eventos con mayor precisión, se podría distribuir los recursos de manera eficiente, evitando que se sobrecarguen las áreas con menos necesidades y enfocando los esfuerzos en las zonas más vulnerables. Esto no solo reduciría el tiempo de respuesta ante emergencias, sino también los costos asociados con el despliegue innecesario de personal y equipos. Además, la mejora en la planificación de los recursos podría llevar a una asignación más eficiente del presupuesto destinado a la seguridad y emergencias, lo cual tendría un impacto positivo en otras áreas del desarrollo urbano.

### **1.3. Objetivos**

#### **1.3.1. Objetivos General**

- Identificar patrones temporales y espaciales en los datos del servicio integrado de seguridad ECU 911 Quevedo mediante la aplicación de algoritmos de clustering para determinar las parroquias de alta peligrosidad.

#### **1.3.2. Objetivos Específicos**

- Analizar la estructura y calidad de los datos proporcionados por el ECU 911 Quevedo para identificar los elementos necesarios para su preprocesamiento

- Aplicar técnicas de preprocesamiento de datos que incluyan la limpieza, normalización y estructuración de la información garantizando la calidad para el análisis de patrones temporales y espaciales.
- Diseñar e implementar algoritmos de agrupamiento adaptándolos a las características de los datos con la finalidad de identificar las parroquias de alta peligrosidad.
- Evaluar la efectividad de los algoritmos y patrones detectados mediante el uso de métricas de desempeño, validando los resultados obtenidos a través de pruebas en diferentes conjuntos de datos.

#### **1.4. Descripción de la Fuente de Datos**

Este proyecto se basa en datos proporcionados por el sistema integrado de seguridad ECU 911, los cuales corresponden a incidentes reportados durante el periodo de enero a diciembre el año 2023, suscitados en la ciudad de Quevedo, dentro del perímetro urbano y rural. Los datos incluyen información temporal y geográfica de los incidentes reportados, que permite identificar patrones de riesgo y áreas de alta peligrosidad. La utilización de estos datos busca generar resultados aplicables para mejorar la seguridad ciudadana mediante el análisis de patrones espaciales y temporales.

## Capítulo 2

### Estado del Arte y la Práctica

#### 2.1. Antecedentes Investigativos

En la investigación titulada: *Detección de patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos* realizada por Chacón Encalada, C. G. en el año 2021 el objetivo general fue identificar patrones en incidentes de emergencia para optimizar la toma de decisiones y reducir costos en el sistema SIS ECU 911 en Ibarra utilizando la metodología de aplicación de técnicas de minería de datos y análisis geoespacial en la que identificaron tendencias en los incidentes y áreas de alta demanda en la ciudad de Ibarra. En el desarrollo de esta investigación se concluyó que la herramienta de minería de datos contribuye a mejorar la gestión de emergencias y operaciones del sistema integrado de seguridad. Dentro de las limitaciones, el estudio enfrenta la dependencia de los datos históricos debido a la baja actualización de información en tiempo real, lo que crea una brecha mediante la combinación de minería de datos, modelos predictivos, y análisis geoespacial. El autor aporta un modelo matemático específico para la localidad de Ibarra, pero puede ser usado en otras ciudades. (Chacon E., 2021)

En la investigación titulada: *Análisis geoespacial de feminicidios y homicidios contra mujeres en Torreón*, realizada por Martínez Narváez, N. A., el objetivo principal del estudio fue identificar patrones geoespaciales en los feminicidios y homicidios para mejorar las estrategias de intervención en Torreón, México, utilizando técnicas de minería de datos y análisis geoespacial. En este trabajo investigativo, se determinó que el análisis de los datos espaciales ayuda a detectar lugares o sectores donde se concentran mayor cantidad de delitos, lo que contribuye a crear estrategias que optimicen los recursos para la prevención del

crimen. Aunque el estudio proporciona un enfoque valioso, enfrenta limitaciones como la pérdida de información crucial en las investigaciones, la falta de datos completos y actualizaciones. El estudio desarrollo puede tener mejoras mediante la integración de múltiples fuentes de datos, incluyendo variables de tipo demográficas y económicas, para enriquecer la caracterización de los puntos críticos y el uso de algoritmos predictivos, como clustering dinámico y redes neuronales, para prever futuros eventos basándose en patrones históricos. El autor aporta un modelo específico y aplicado a la ciudad de Torreón, podría ser adaptable a otras ciudades con características similares de eventos delictivos. (Martinez N., 2023)

En la investigación titulada: *Aplicaciones de algoritmos de clustering en inteligencia artificial*, se llevó a cabo por un equipo de expertos en IA, y se centró en explorar cómo los algoritmos de clustering pueden aplicarse en distintos campos, como la gestión de emergencias, para analizar grandes cantidades de datos. Para ello, se utilizaron tanto algoritmos jerárquicos como no jerárquicos, con el objetivo de segmentar eventos de emergencia y priorizar las respuestas. Sin embargo, el estudio no se enfoca específicamente en emergencias geoespaciales, lo que limita su aplicación directa a situaciones como desastres naturales o crisis. Aunque se describen ambos tipos de algoritmos, no se realiza una comparación de su efectividad para problemas específicos, lo que podría haberse abordado a través de un benchmarking entre ambos enfoques. En esta investigación se concluyó que los algoritmos de clustering mejoran la eficiencia en la gestión de los datos de emergencias, pero se recomienda realizar comparaciones más detalladas entre diferentes métodos y adaptarlos al contexto geográfico. Este enfoque aplicado en la investigación es valioso, debido a que este permite realizar un análisis de manera más precisa y técnica de los datos geoespaciales y la asignación de recursos en tiempo real. (Universitat Internacional Valenciana., 2022)

En la investigación titulada: *Análisis espacio-temporal de robos a personas en el Centro Histórico de Cuenca–Ecuador: Comparación entre denuncias judiciales y reportes de emergencia para el periodo 2019*, realizada por López, E. B., & Inga, O. D., tuvo como finalidad realizar la evaluación de los patrones detectables en los robos efectuados dentro del Centro Histórico de Cuenca mediante la comparación de denuncias judiciales y reportes de emergencia. La metodología que utilizaron se basó básicamente en el análisis de los datos de los incidentes y la aplicación de técnicas de visualización geoespacial finalizando con unas discrepancias significativas entre los reportes de emergencia y las denuncias judiciales de estos eventos.

En esta investigación se concluyó que mejorar la recopilación de datos puede aumentar la eficacia en la respuesta a los robos. Los autores aportan un enfoque comparativo que resalta la necesidad de mejorar la comunicación entre las autoridades y la población. (López, E. B., & Inga, O. D., 2022)

En la investigación titulada: *Visualización de datos Espacio-Temporales de predicción de Eventos de Emergencia en el Ecuador*, realizada por Villamar Barros, H. F. en el año 2023, el objetivo general incluyó desarrollar un sistema de visualización de datos espacio-temporales para predecir eventos de emergencia en Ecuador. La metodología que empleó el investigador permitió realizar el análisis de datos históricos y la técnica de visualización avanzada, obteniendo como resultado la detección de patrones de riesgo utilizables en la asignación de recursos de las entidades gubernamentales para la atención de eventos de seguridad ciudadana. En esta investigación se concluyó que la visualización de datos mejora la toma de decisiones en situaciones críticas. El autor aporta un enfoque innovador para la gestión de emergencias basado en datos visuales. (Villamar H., 2023)

En la investigación titulada: *Análisis estadístico espacio-temporal de los delitos causantes de la inseguridad ciudadana en la provincia de Chimborazo del 2015 al 2019*,



realizada por Valdez Huaraca, F. P. durante el año 2021, se planteó en su estudio evaluar la incidencia de los delitos en la provincia de Chimborazo y su relación con la inseguridad ciudadana. La metodología utilizada incluyó técnicas estadísticas y análisis geoespacial sobre datos de delitos, identificándose patrones temporales y espaciales de delincuencia. En esta investigación se concluyó que el análisis permite comprender mejor la inseguridad en Chimborazo. El autor aporta una base estadística que puede ser utilizada para diseñar estrategias de intervención. (Valdez F., 2021)

En la investigación titulada: *Seguridad electrónica aplicada a centrales de emergencia urbana 911, Piura*, realizada por Chiyon Arrunátegui, R. E. en el año 2021, se realizó con el objetivo de evaluar la implementación de sistemas de seguridad electrónica en las centrales del servicio de seguridad 911 en la ciudad de Piura. La metodología que el investigador empleó consistió en un enfoque de tipo descriptivo mediante el análisis de casos y entrevistas realizadas al personal del área de seguridad, identificando deficiencias en los sistemas actuales y proponiendo recomendaciones para su mejora. Mediante el estudio de definió que la seguridad electrónica ayuda de manera considerable a incrementar la eficiencia y eficacia ante una respuesta a emergencias que se susciten en la ciudad. Sin embargo, la investigación enfrenta limitaciones como la implementación parcial de la tecnología en la entidad y la dependencia que esta tiene de la infraestructura que no es la adecuada.

Adicional, el estudio también subraya la importancia de capacitar de forma constante al personal de seguridad y de mejorar la infraestructura tecnológica de la entidad, y esta permita asegurar que los sistemas de seguridad electrónica funcionen de manera eficiente ante las emergencias que se suscitan diariamente. Estos enfoques permitirán superar las barreras de implementación y maximizar el impacto de las mejoras propuestas. (Chiyon R., 2021)

## **2.2. Fundamentación Legal**

### **2.2.1. Legislación Internacional**

- Convención de las Naciones Unidas sobre el Derecho Penal de 1988: Este documento establece principios básicos para la cooperación internacional en la lucha contra el crimen, incluyendo la seguridad ciudadana y la prevención de delitos.
- Declaración de los Derechos Humanos: Artículos que abogan por el derecho a la seguridad y protección de las personas, que fundamentan la importancia de un sistema de seguridad efectivo.

### **2.2.2. Legislación Nacional**

- Código Orgánico Integral Penal (COIP):

Artículo 1: Tiene como objetivo asegurar la seguridad de la ciudadanía y promover una convivencia pacífica.

Artículo 2: En materia penal se refiere a la definición de las competencias que tienen los organismos encargados de la seguridad pública, garantizando el principio de legalidad y protección de los derechos humanos.

- Ley Orgánica de Seguridad Ciudadana y del Estado:

Artículo 3: En este apartado se establecen los principios de la seguridad ciudadana, enfatizando la prevención del delito.

Artículo 5: Consiste en los mecanismos para el uso y aplicación de las tecnologías en la gestión de la seguridad

- Ley de Protección de Datos Personales:

Artículo 4: Establece el marco para la recolección y tratamiento de datos personales, relevante para el manejo de información en el sistema ECU 911.

Adicional, el artículo 66 de la Constitución de la República, en su numeral 19, reconoce y garantiza el derecho que los ciudadanos tienen sobre la protección de los datos de carácter personal, lo que, además, incluye el acceso y la decisión sobre la información de este carácter, lo que derivaba a su correspondiente protección.

## **2.3. Fundamentación Teórica**

### **2.3.1. Seguridad Ciudadana**

La seguridad ciudadana abarca todas las acciones, políticas y medidas que buscan proteger a las personas frente a los riesgos que afectan su bienestar. Esto incluye prevenir y reducir los delitos, proteger contra la violencia y asegurar que los derechos humanos sean respetados tanto en espacios públicos como privados. La seguridad ciudadana no es solo responsabilidad de las fuerzas del orden, sino también de la comunidad, ya que su participación, el bienestar social y las políticas públicas que fomentan la unidad son clave para lograr un entorno más seguro y permitan una convivencia pacífica dentro de la sociedad (Zapata, S., Pérez, M., & Gutiérrez., 2018)

#### **2.3.1.1. Sistema Integrado de Seguridad ECU 911**

Conocido como Ecu 911 es la entidad encargada de gestionar las emergencias dentro del territorio ecuatoriano, esta entidad pública coordina y gestiona las llamadas de emergencias que los ciudadanos realizan a diario, facilitando una respuesta rápida ante los distintos incidentes que se suscitan. Su objetivo es dar respuesta a las solicitudes de auxilio recibidas, identificar el tipo de incidente y derivar el recurso necesario respectivo de acuerdo al evento presentado, coordinando con entidades como la policía, bomberos y unidades de salud. Además, ofrece servicios de monitoreo de cámaras de seguridad y apoyo en la gestión de crisis a nivel nacional.

El servicio busca optimizar los tiempos de respuesta y mejorar la coordinación interinstitucional para resolver emergencias de manera eficiente (Vega, L, 2020)

**Figura 1.**

*Modelo de gestión sistema integrado de seguridad ECU 911*



Nota: Grafico obtenido del Ecu 911

### 2.3.1.2. Problemas Actuales de Seguridad en Quevedo

Quevedo, una ciudad intermedia de Ecuador, enfrenta varios retos en cuanto a seguridad, destacándose el incremento de delitos como robos, violencia doméstica y narcotráfico. Los recursos que posee actualmente la policía son deficientes, y esto limita la cobertura que pueda dar. El uso de cámaras de vigilancia es mínimo y la sensación de impunidad en el territorio agravan la situación. Estas dificultades tienen su origen, en parte, en crecimiento poblacional acelerado sin una planificación adecuada, así como en la falta de infraestructura en áreas más vulnerables. (Sánchez, M., & Torres, J., 2019)

### 2.3.1.3. Estrategia de Gestión del ECU 911

El ECU 911 ha implementado una serie de estrategias en Quevedo, tales como la expansión de la red de cámaras de seguridad, la mejora en la formación de su personal y la optimización de los tiempos de respuesta en situaciones de emergencia. Además, se ha promovido la colaboración con otras instituciones locales para coordinar esfuerzos y mejorar la efectividad en la respuesta a incidentes. El objetivo de estas medidas es reducir los tiempos de reacción y fortalecer la confianza de la ciudadanía en los servicios de emergencia. (Ramírez, D., Mora, G., & Peña, J., 2020)

**Figura 2.**

*Estrategia de gestión*



Nota: Gráfico obtenido del Ecu 911

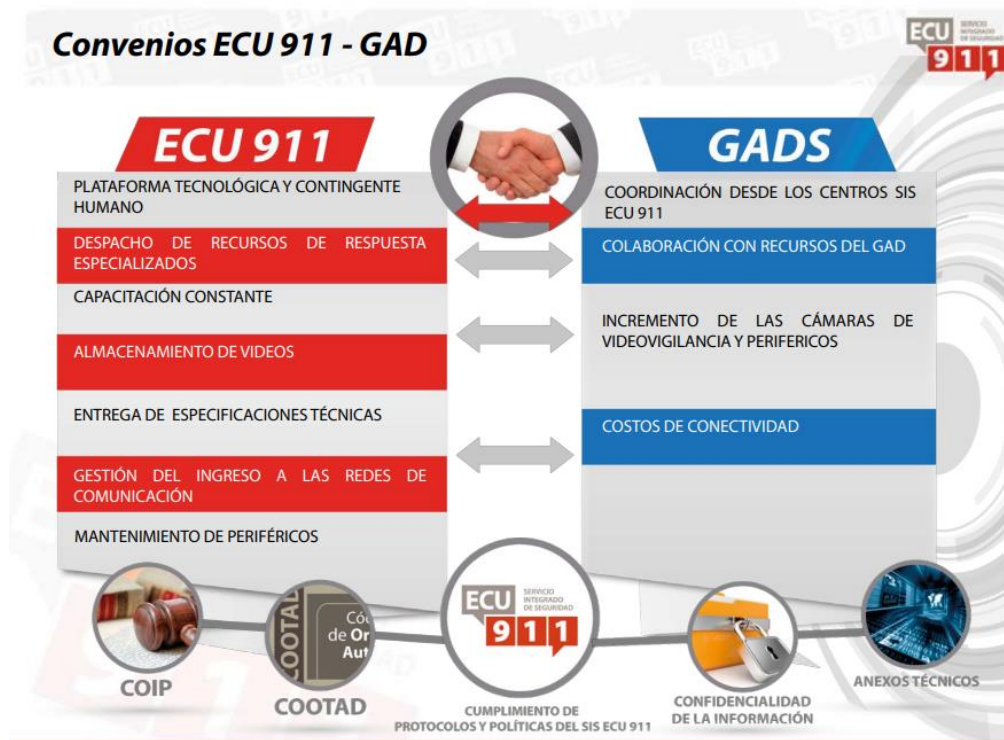
### 2.3.1.4. Coordinación Interinstitucional en la Gestión de Emergencias

La coordinación que se tena entre instituciones es de vital importancia para tener una respuesta inmediata en situaciones de emergencia o catástrofes. En la ciudad de Quevedo

existe cooperación entre las diferentes entidades como el ECU 911, la policía, los bomberos y las unidades de salud lo que resulta clave para asegurar respuestas rápidas y eficientes. Sin embargo, las dificultades en la comunicación, los pocos recursos, la planificación y la falta de gestión representan obstáculos que afectan la eficiencia y eficacia de esta cooperación. El desarrollar mejora en los protocolos de colaboración y fortalecimiento de las capacidades locales son pasos fundamentales para optimizar la gestión de emergencias dentro del territorio. (González, S., & López, M., 2019)

**Figura 3.**

*Coordinación de Interinstitucional de emergencias*



### 2.3.2. Patrones espaciales y temporales

Los patrones espaciales se refieren a la distribución geográfica de eventos o fenómenos en un área específica. Estos pueden manifestarse de manera regular, aleatoria o concentrada en puntos particulares, y su análisis permite identificar zonas con mayor o

menor actividad, lo que es clave en ámbitos como la seguridad pública. Los patrones temporales estudian el desarrollo de los eventos en función del tiempo, describiendo como estos se distribuyen permitiendo identificar tendencias como la estacionalidad, los picos de actividad, la concentración horaria, el tiempo de baja incidencia. Comprender y detectar este tipo de patrones es crucial para la elaboración de planificaciones, la asignación de recursos y la predicción de posibles eventos que podrían ocurrir. (Tobón, A., Silva, J., & Pérez, M., 2018)

El análisis de patrones espaciales es clave para garantizar la seguridad ciudadana, ya que ayuda a identificar zonas de alto riesgo, también conocidas como 'puntos calientes', donde los delitos y actividades ilícitas tienden a centrarse. Esta información es fundamental para distribuir los recursos de seguridad de manera más eficiente, permitiendo una intervención más precisa y preventiva. Además, comprender cómo se distribuyen los delitos en un área geográfica específica facilita la creación e implementación de políticas públicas que respondan de manera más eficiente a las necesidades de cada zona, mejoren los índices de medición y contribuya a la aplicación de políticas públicas de seguridad. El análisis espacial permite, por lo tanto, una planificación estratégica, permitiendo optimizar el uso de los recursos y reducir los índices de criminalidad en puntos actualmente considerados como críticos. (Braga, A., Welsh, B., & Schnell, C., 2019)

### **2.3.2.1. Clasificación de patrones espaciales**

Existen varias formas de clasificar los patrones espaciales, dependiendo de su distribución y comportamiento. Las principales categorías incluyen según Andresen y Malleso en 2023:

- **Patrones de Densidad:** se refieren a la concentración de eventos en áreas específicas, lo que indica una mayor ocurrencia de fenómenos en esas zonas.

En el contexto de la seguridad, esto puede reflejar una mayor frecuencia de delitos en determinados puntos geográficos, lo cual ayuda a identificar áreas con mayor riesgo y a planificar intervenciones focalizadas. Este tipo de análisis es fundamental para dirigir los recursos y esfuerzos de manera más eficiente hacia las áreas de mayor incidencia delictiva.

- **Patrones de distribución:** son aquellos en la que los eventos se dispersan sobre un área determinada, sea de forma aleatoria, uniforme o agrupada. Este tipo de análisis contribuye en la comprensión de tendencias de acuerdo a la forma en cómo se agrupan los delitos dentro de una región. Si un patrón es de forma aleatoria indica que no tiene una tendencia predefinida, y que el evento no tiene un patrón predecible, mientras que los patrones uniformes o agrupados reflejan concentraciones de actividades en áreas específicas, lo que puede señalar puntos críticos que requieren atención especial en términos de seguridad
- **Patrones de proximidad:** Estos patrones analizan la relación espacial entre los eventos, mediante la evaluación espacial de los eventos diagnostican si estos se agrupan en áreas cercanas o se distribuyen de manera más dispersa. Este enfoque es útil para identificar áreas con alta concentración de incidentes o para observar cómo los fenómenos se distribuyen a lo largo de un territorio.  
(Andresen, M. A., & Malleson, N., 2023)

### **2.3.2.2. Métodos para identificar patrones delictivos**

- **Métodos Estadísticos:** Mediante el uso de métodos matemáticos se identifica y detecta patrones, esto se realiza a través del uso de herramientas como el análisis de series temporales, mapas de calor y modelos de regresión para examinar la distribución y las tendencias de los eventos. Estas técnicas permiten tener una



mejor comprensión de los fenómenos, facilitando la predicción y la toma de decisiones basadas en datos cuantitativos. (Gutiérrez, C., & Sánchez, M., 2019)

- **Métodos algorítmicos:** Emplea el uso de algoritmos de aprendizaje automático y de las técnicas de minería de datos, lo que permite descubrir patrones subyacentes en grandes bases de datos relacionadas con diferentes eventos o delitos. Estos métodos permiten detectar tendencias ocultas en la información, como patrones de criminalidad, comportamientos recurrentes o áreas de alta incidencia, sin la intervención directa del analista. El uso de estos enfoques da como resultado la predicción de futuros incidentes, lo que conlleva a optimizar la asignación de recursos de los diferentes entes y mejora a nivel jerárquico la toma de decisiones en la gestión de la seguridad. (Gutiérrez, C., & Sánchez, M., 2019)
- **Métodos Visuales:** Estos métodos se enfocan en representar visualmente los datos mediante herramientas como mapas de calor, diagramas de dispersión y otras formas de visualización, cuyo objetivo es facilitar la interpretación de patrones complejos mediante información visual. Estas representaciones gráficas permiten a los analistas identificar rápidamente tendencias y concentraciones, mejorando la comprensión y toma de decisiones sobre la distribución de recursos y las asignaciones de atención. Además, estas visualizaciones permiten una comunicación directa, una interpretación clara entre las diferentes entidades y obtener resultados positivos ante las acciones planteadas por las autoridades de seguridad o los responsables de la toma de decisiones en cada Gobierno Autónomo Descentralizado. (Gutiérrez, C., & Sánchez, M., 2019)

### 2.3.2.3. Factores que Influyen en los Patrones Delictivos

**Tabla 1.**

*Factores que influyen en los patrones delictivo*

Factor	Descripción
Factores Demográficos	La edad, el género y el nivel educativo de los individuos influyen en la probabilidad de cometer delitos.
Factores Socioeconómicos	La pobreza, el desempleo y la desigualdad social son factores que aumentan la incidencia delictiva.
Factores Ambientales	El diseño urbano y la presencia de zonas vulnerables influyen en la ocurrencia de delitos.
Factores Temporales	Los delitos varían según la temporada, como los robos durante las vacaciones o ciertos horarios del día.

Nota: (Muñoz, V., & Ramos, J., 2019)

### 2.3.3. Herramientas y Métodos para el Análisis de Patrones

#### 2.3.3.1. Sistema de Información Geográfico (SIG) en la Seguridad Pública

Los Sistemas de Información Geográfica (SIG) son herramientas utilizadas por servicio integrado de seguridad ECU 911 para analizar fenómenos espaciales, como los patrones delictivos desarrollados en el territorio ecuatoriano mediante la recopilación, análisis y visualización de los datos geoespaciales en el sistema mencionado. Teniendo como resultado la identificación de zonas con alta incidencia delictiva y detección de patrones espaciales.

La visualización de los datos de forma esquemática permite tener una perspectiva más clara de las necesidades que tiene cada provincia, cada cantón y cada parroquia. Lo que mejora la planificación de cada una de las entidades inmersas en el bienestar ciudadano, mediante la implementación de estrategias de seguridad focalizadas más precisas y efectivas. (Esri., 2021)

**Tabla 2.**

*Aspectos del Sistemas de Información Geográfica*

Aspecto	Descripción
Aplicaciones de SIG en el Contexto Urbano	Los SIG permiten visualizar y analizar la distribución geográfica de delitos en zonas urbanas, identificando áreas de alta incidencia y patrones delictivos, lo que facilita una planificación más efectiva de los recursos de seguridad.
Ventajas del Análisis Geoespacial	El análisis geoespacial permite integrar datos sobre delitos, infraestructura, uso del suelo y factores socioeconómicos, proporcionando un enfoque integral para la prevención del crimen.
Limitaciones de los SIG en la Seguridad Urbana	A pesar de sus ventajas, los SIG requieren datos precisos y actualizados, lo que puede ser costoso. Además, integrar información en tiempo real puede ser un desafío técnico.
Casos Reales de Aplicación de SIG	Varias ciudades, como Los Ángeles, utilizan SIG para mapear delitos y mejorar la distribución de recursos, optimizando la respuesta ante incidentes.

Nota: Esri, 2021

### **2.3.3.2. Mapas de Calor**

Los mapas de calor son representaciones gráficas que concentra eventos de diferentes tipos dentro de una referencia o región geográfica. Mediante la utilizan un rango de colores, que va desde los colores fríos hasta los cálidos, de manera visual permite identificas áreas con mayor frecuencia de un fenómeno específico. Esta herramienta es útil para identificar patrones espaciales y enfocar los recursos en zonas con mayor incidencia, mejorando la eficiencia en la planificación y las intervenciones de seguridad, permitiendo aplicar medidas correctivas en las áreas que sean consideradas de alto riesgo.

Los mapas de calor deben interpretarse con cuidado, ya que las zonas con mayor actividad delictiva no siempre son las únicas que necesitan atención. Es importante considerar otros factores, como la densidad poblacional, para tener una comprensión más completa y precisa de los patrones, lo que permitirá priorizar de forma efectiva las áreas que

realmente requieren intervención, unir estos datos con otro tipo de análisis y corroborar los resultados obtenidos.

Las ciudades utilizar los mapas de calor para asignar las patrullas de forma eficiente. En Chicago, estos mapas se actualizan en tiempo real y permiten predecir y prevenir delitos en áreas específicas. Herramientas como ArcGIS y QGIS permiten crear mapas de calor efectivos a partir de bases de datos espaciales y temporales. (Gibson, C., 2020)

**Figura 4.**

*Mapa de calor*



Nota: [https://www.fii.gob.ve/mapas-de-calor-en-los-sistemas-de-informacion-geografica/\(2024\)](https://www.fii.gob.ve/mapas-de-calor-en-los-sistemas-de-informacion-geografica/(2024))

**2.3.3.3. Técnicas de Minería de Datos para la Seguridad**

La minería de datos en seguridad pública consiste en aplicar métodos estadísticos y algoritmos para identificar patrones ocultos en grandes conjuntos de datos criminales. Este proceso incluye técnicas como el análisis de agrupamientos, la detección de anomalías y la

clasificación de datos, lo cual permite descubrir tendencias que podrían no ser evidentes a simple vista, facilitando la toma de decisiones en la prevención y respuesta a delitos.

Mediante la minería de datos se identifican patrones delictivos a través del estudio y análisis de grandes bases de datos obtenidas de crímenes, en la que se aplica técnicas como el clustering, detección de anomalías, y redes neuronales. (Pedregosa, F., et al., 2020)

- Algoritmos de Clustering

Los algoritmos de clustering son herramientas fundamentales en la minería de datos, consiste en la agrupación de elementos similares dentro de un conjunto de datos de una o varias categorías. Estos algoritmos tienen como finalidad identificar estructuras subyacentes en los datos, dividiéndolos en grupos o mejor definidos clústeres, basados en medidas de similitud o proximidad, como la distancia euclidiana.

### **K-Means**

El algoritmo minimiza la suma de las distancias cuadradas de cada punto  $x_i$  al centroide de su clúster  $C_k$ :

$$J = \sum_{k=1}^k \sum_{x_i \in C_k} \|x_i - \mu_k\|^2$$

donde  $k$  es el número de clústeres,  $\mu_k$  es el centroide del clúster  $C_k$ , y  $\|\cdot\|$  representa la distancia euclidiana. Este enfoque iterativo converge al mínimo local. Este método particional divide los datos en un número predefinido de clústeres, minimizando la suma de las distancias cuadradas entre los puntos y los centroides del grupo al que pertenecen. Es eficaz para datos bien separados, pero sensible a los valores atípicos donde incide mayormente los datos que generen ruido. (García, J., Martínez, P., & López, R., 2023)

## DBSCAN

Basado en densidad, este algoritmo identifica clústeres de cualquier forma al considerar zonas densamente pobladas y separarlas de áreas de menor densidad. Es adecuado para datos con ruido, pero su rendimiento depende de los parámetros seleccionados (radio y número mínimo de puntos).

$\varepsilon$  = radio de búsqueda

*MinPts* = indica el número mínimo de puntos en una región. Un punto *ppp* es central si cumple:

$$|N_\varepsilon(p)| \geq \text{MinPts}$$

donde  $|N_\varepsilon(p)|$  es el vecindario dentro de  $\varepsilon$ .

Este método agrupa datos contiguos en densidad y marca el resto como ruido. (García, J., Martínez, P., & López, R., 2023)

## Clustering Jerárquico

Construye una jerarquía de clústeres mediante divisiones sucesivas o de combinaciones (enfoque aglomerativo), lo que permite que el investigador pueda analizar las relaciones en diferentes niveles de granularidad. Forma una dendrograma utilizando la distancia mínima, máxima o promedio entre clústeres, lo que permite visualizar como los datos se agrupan de forma jerárquica. A nivel matemático, se optimiza:

$$D(C_i, C_j) = \min \|x_i - x_j\|$$

Dónde,

$C_i$  y  $C_j$  son clústeres.

## HDBSCAN

En español, sus siglas indica agrupamiento espacial jerárquico basado en densidad de aplicaciones con ruido, es un algoritmo de agrupamiento que combina lo mejor del algoritmo DBSCAN y el agrupamiento jerárquico. A diferencia de métodos como K-means, HDBSCAN no necesita que se especifique el número de grupos con antelación, lo que, si sucede en K-means, lo que lo convierte en una excelente opción para identificar patrones en datos con densidades variables. Este algoritmo organiza todos los datos en una jerarquía de clústeres y selecciona la estructura más adecuada al calcular la densidad local de los puntos.

Utiliza la distancia que existe entre cada punto y sus vecinos cercanos para formar grupos, facilitando la detección de patrones en datos complejos y ruidosos, como aquellos utilizados en el análisis. (Díaz, D., & Rodríguez, M., 2023)

- Detección de Anomalías y Outliers

La detección de anomalías se caracteriza por identificar datos que se desvían significativamente de los patrones generales de un conjunto. Este proceso se fundamenta principalmente en el uso de herramientas estadísticas y algoritmos avanzados, diseñados para medir de forma significativa desviaciones existentes y reconocer patrones inesperados dentro de grandes volúmenes de datos. (Liu, F. T., Ting, K. M., & Zhou, Z.-H. , (2020). )

Métodos estadísticos:

$$P(x_i) < \delta$$

donde  $\delta$  es un umbral definido

Distancia euclidiana

$$D(x_i, \mu) = \|x_i - \mu\|$$

Dónde, donde  $x_i$  es un punto y  $\mu$  el promedio de los datos

Modelos de Machine Learning:

Isolation Forest: Aísla datos atípicos creando particiones aleatorias.

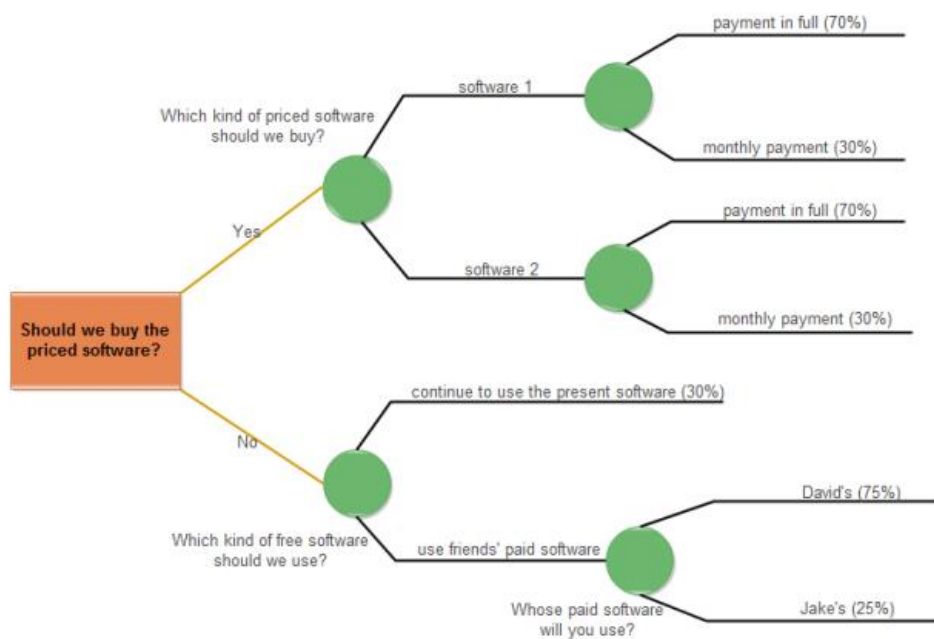
Autoencoders: Reconstruyen datos normales; los errores altos indican anomalías.

- Árboles de Decisión:

Los Árboles de decisión son herramientas de análisis predictivo que permiten clasificar resultados mediante un sistema jerárquico de decisiones. En el campo de la seguridad, estos modelos se utilizan para identificar patrones delictivos a partir de datos históricos. Matemáticamente, consiste en dividir el espacio de los datos en subconjuntos más pequeños de manera iterativa, empleando métricas para medir la eficacia de cada partición. (Safavian, S. R., & Landgrebe, D., 2020)

**Figura 5.**

*Árbol de decisión*



Nota: Safavian, S. R., & Landgrebe, D. (2020)



## Capítulo 3

### Diseño Metodológico

#### 3.1. Enfoque de la Investigación

El enfoque de la investigación se orienta a un análisis de tipo cuantitativo mediante el análisis estadístico de los datos otorgados por el servicio integral de seguridad ECU 911, con un diseño exploratorio, utilizando métodos de *clustering* para identificar patrones de tipo espaciales y temporales. El análisis se centra en la recolección y evaluación de información geoespacial y temporal proporcionada por la entidad Matriz, facilitando los datos específicos de la ciudad de Quevedo, con el objetivo de identificar áreas de alta peligrosidad. La investigación tiene un carácter aplicado, buscando optimizar la distribución de recursos y la efectividad en la atención de los eventos de seguridad.

#### 3.2. Diseño de la Investigación

Se aplica un diseño no experimental de tipo correlacional, la finalidad es explorar cómo se relacionan las variables geoespaciales, como la ubicación y distribución de los delitos, con las tasas de criminalidad en las parroquias, sin intervenir directamente en estas variables. Los datos son obtenidos mediante trámite formal en la página Contacto Ciudadano entidad servicio ECU 911 matriz, garantizando su veracidad y relevancia. La recolección y el análisis de estos datos se realizarán de manera sistemática, con el fin de identificar patrones y detectar las zonas más vulnerables. Además, se busca optimizar la asignación de recursos para mejorar la efectividad de las intervenciones de seguridad en las parroquias identificadas y en los tiempos determinados.

#### 3.3. Tipo de Investigación

El tipo de la investigación es aplicado y exploratorio. El estudio es aplicado ya que está orientado a solucionar un problema práctico que se desarrolla en la ciudad de Quevedo, mediante la detección de zonas de alto riesgo en el tema de seguridad ciudadana y a su vez

permita la optimización de recursos a través del análisis de datos espaciales y temporales provenientes del ECU 911

Es exploratoria dado que tiene como objetivo identificar patrones y relaciones que previamente son desconocidas entre las variables geoespaciales y las tasas delictivas actuales en las zonas rurales y urbanas de la ciudad, sin una hipótesis previa definida.

### **3.4. Nivel de Investigación**

El estudio tiene un nivel de investigación exploratorio y descriptivo. Es exploratorio porque busca identificar patrones y relaciones que previamente no son conocidas en los datos proporcionados por el ente ECU 911. También es descriptivo, ya que se centra en examinar detalladamente la distribución y características de los patrones delictivos en Quevedo, a partir de los datos de seguridad y su correlación geoespacial durante el año 2023.

### **3.5. Técnicas e Instrumentos de Recolección de Datos**

Para este estudio, se emplea técnicas cuantitativas de recolección de datos, mediante un proceso formal, vía informática se genera ticket SIS\_ECU\_911-SOL-2024-010795 en página Contacto Ciudadano Digital en el que se obtiene la información proporcionada por el servicio integrado de seguridad ECU 911. Esta información es entregada con la firma de un acta de confidencialidad y copia de documento de identidad, lo que asegura que los datos se empleen exclusivamente para los fines de la investigación. Posteriormente, se emplean Python (Google colab) herramienta estadística y de análisis geoespacial, para procesar los patrones espaciales y temporales y detectar las áreas de alta peligrosidad en Quevedo

### **3.6. Técnicas para el Procesamiento e Interpretación de Datos**

En este estudio, se utilizan algoritmos de clustering para analizar los incidentes delictivos reportados por el ECU 911, aprovechando herramientas de Python. Estos métodos fueron elegidos porque permiten identificar patrones espaciales y temporales en grandes

volúmenes de datos, lo que ayuda a detectar las zonas con mayor índice de criminalidad en Quevedo.

En la realización de la segmentación de los datos, se aplicaron los algoritmos K-Means, DBSCAN y HDBSCAN, utilizando bibliotecas especializadas como scikit-learn y HDBSCAN en Google Colab mediante el uso de código de Python. La selección del algoritmo más adecuado se realizó mediante métricas de evaluación, destacando el índice de Silhouette, que mide la coherencia y calidad de los grupos formados y adicional a este el índice DBI.

Este enfoque permite analizar de manera objetiva la distribución de los delitos, asegurando que los patrones detectados definan la realidad delictiva de las parroquias en todo el perímetro de la ciudad. Además, se aplicaron medidas para controlar la variabilidad en los registros, aplicando filtros que eliminan datos atípicos y mejoren la precisión del análisis. De esta manera, los resultados obtenidos pueden servir como una herramienta clave para la toma de decisiones en materia de seguridad y asignación de recursos.

### **3.7. Población y Muestra**

#### **3.7.1. Población**

La población de estudio está compuesta por los datos generados de los incidentes delictivos registrados por el servicio ECU 911 en la ciudad de Quevedo. Esta población incluye los reportes de delitos en diferentes parroquias de la ciudad, la información dotada por el sistema integrado de seguridad consta desde enero del 2023 a Julio del 2024, teniendo un total de 68551 incidentes reportados.

#### **3.7.2. Tamaño de la Muestra**

En este estudio, la población de análisis está compuesta por los registros de incidentes delictivos reportados al servicio ECU 911 durante un período de un año completo, desde enero hasta diciembre de 2023 en la ciudad de Quevedo. La elección de un solo año

de datos responde a la necesidad de contar con información actualizada, permitiendo identificar patrones y tendencias recientes en la actividad delictiva, lo cual es clave la toma de decisiones estratégicas. Además, analizar un período más extenso podría generar distorsiones debido a cambios en las políticas de seguridad, variaciones demográficas o eventos específicos que alteran las dinámicas delictivas.

El tamaño de la muestra corresponde a la totalidad de los incidentes registrados en el periodo enero a diciembre del año 2023, sumando un total de 45.674 casos. Esta decisión garantiza que el análisis refleje con precisión las fluctuaciones en la distribución geoespacial y temporal de los delitos, asegurando una representación clara de las zonas de mayor peligrosidad en la ciudad. además, al utilizar los datos de un año completo minimiza el sesgo muestral y se optimiza la aplicación de algoritmos de clustering, permitiendo extraer patrones más representativos sin necesidad de ajustes adicionales en el tamaño de la muestra.

Para manejar la variabilidad en los registros de incidentes, se tomarán en cuenta factores como la estandarización de las fuentes de datos, la eliminación de posibles duplicados y la clasificación de los eventos según el tipo de delito y su ubicación. De esta forma, se asegura que los resultados sean consistentes y puedan utilizarse para diseñar estrategias de seguridad basadas en datos reales.

Para controlar la variabilidad en los registros de incidentes dentro del estudio, se implementaron las siguientes estrategias:

1. Estandarización de Fuentes de Datos:

- Se utilizaron únicamente registros oficiales del servicio de seguridad integral ECU 911, asegurando que la información proviene de una fuente fidedigna.
- Se verificó que los datos siguieran un formato unificado en cuanto a fecha, ubicación y tipo de incidente.

## 2. Depuración y Validación de Datos:

- Se eliminaron registros duplicados y aquellos con información incompleta o errónea para evitar sesgos en el análisis.

## 3. Segmentación por Tipología y Ubicación:

- Se agruparon los incidentes por tipo delictivo y zona geográfica, permitiendo analizar patrones específicos sin que la variabilidad entre categorías afecte los resultados generales.
- Esto permite detectar si algunos delitos muestran variaciones más marcadas en ciertas zonas o en determinados períodos.

## Capítulo 4

### Análisis y Discusión de los Resultados

#### 4.1. Análisis Descriptivo de los Resultados

##### 4.1.1. Organización de los Datos

La ciudad de Quevedo pertenece a la provincia de Los Ríos y en el contexto del servicio de seguridad integrado ECU 911, la ciudad está dividida en circuitos y subcircuitos para mejorar la eficiencia en la respuesta ante incidentes (ECU 911, 2023). Según la organización del ECU 911, los circuitos y subcircuitos representan áreas específicas de la ciudad donde se monitorean y gestionan las emergencias, a la cual el GAD municipal y gobierno central asigna recursos para el desarrollo cotidiano de sus actividades.

Cada circuito está compuesto por un número de subcircuitos definidos dependiendo de la dimensión geográfica lo que permite tener un mayor control de las distintas parroquias en las que se encuentra dividida la ciudad. Esta división geográfica y administrativa es fundamental para analizar los patrones temporales y espaciales de los incidentes que se han reportado durante todo el año, ya que permite asociar las parroquias y los tipos de incidentes con los circuitos y subcircuitos correspondientes que se suscitan durante el transcurso del día.

Esto puede ayudar a identificar áreas de alto riesgo o de mayor frecuencia de incidentes, lo que es crucial para la seguridad y la planificación de intervenciones de manera zonificada, tanto en la zona rural como urbana de la ciudad.

En la Tabla 3 se visualizan los circuitos y subcircuitos en los que se encuentra dividida la ciudad, teniendo a la parroquia Galo plaza, Divino niño y 7 de octubre con 3 subcircuitos debida a su dimensionamiento geográfico.

**Tabla 3.***Código de los circuitos y subcircuitos del cantón Quevedo*

Cod. Circuito	Nombre del circuito	No. de Subcircuitos	Cod. Subcircuitos	Nombre Subcircuitos
12d03c01	La Esperanza	1	12d03c01s01	La Esperanza 1
12d03c02	Control (Centro)	2	12d03c02s01	Control 1
		2	12d03c02s02	Control 2
12d03c03	La Valdramina	1	12d03c03s01	La Valdramina 1
12d03c04	La Merced	2	12d03c04s01	La Merced 1
		2	12d03c04s02	La Merced 2
12d03c05	Galo Plaza	3	12d03c05s01	Galo Plaza 1
		3	12d03c05s02	Galo Plaza 2
		3	12d03c05s03	Galo Plaza 3
12d03c06	Divino Niño	3	12d03c06s01	Divino Niño 1
		3	12d03c06s02	Divino Niño 2
		3	12d03c06s03	Divino Niño 3
12d03c07	7 de Octubre	3	12d03c07s01	7 de Octubre 1
		3	12d03c07s02	7 de Octubre 2
		3	12d03c07s03	7 de Octubre 3
12d03c08	Guayacán	1	12d03c08s01	Guayacán 1
12d03c09	San Carlos	1	12d03c09s01	San Carlos 1

**4.1.1.1. Incidentes por Circuito**

Durante el año 2023, en la ciudad de Quevedo existieron 45674 incidentes reportados mediante llamada al Ecu 911, los cuales se registraron de la siguiente manera (Ver tabla 4).

**Tabla 4.***Incidente por Circuito*

Parroquia	Cod_Circuito	Número de Incidentes
7 de Octubre	12d03c07	14389
Galo Plaza	12d03c05	9798
Divino Niño	12d03c06	7258
Control	12d03c02	2970
San Carlos	12d03c09	2909
La Merced	12d03c04	2623
La Valdramina	12d03c03	2556
La Esperanza	12d03c01	2219
Guayacán	12d03c08	952

La tabla presenta una distribución de los incidentes registrados en distintas parroquias de Quevedo, organizadas por circuito y con la cantidad de incidentes reportados en cada una. Este tipo de información es esencial para comprender la dinámica de los incidentes en la ciudad y para identificar patrones de mayor riesgo

Las parroquias 7 de Octubre y Galo Plaza son las parroquias con los mayores registros de incidentes, con 14.389 y 9.798 incidentes, respectivamente. Esta alta incidencia podría estar relacionada con su mayor densidad poblacional, su relevancia comercial o su ubicación geográfica central dentro de Quevedo. El hecho de que estas parroquias concentren una parte significativa de los incidentes podría indicar una mayor exposición a situaciones de riesgo, lo que justificaría una atención prioritaria en la distribución de recursos para la seguridad. Guayacán es la parroquia con el menor número de incidentes, con solo 952. Esto podría reflejar una menor densidad poblacional o una menor actividad social o económica en esa área.



#### 4.1.1.2. Incidentes por Subcircuito

Tabla 5.

*Incidente por Subcircuito*

Parroquia	Cod_Subcircuito	Número de Incidentes
7 de Octubre	12d03c07s02	7004
7 de Octubre	12d03c07s01	6419
Divino Niño	12d03c06s02	5907
Galo Plaza	12d03c05s03	4084
Galo Plaza	12d03c05s01	3868
San Carlos	12d03c09s01	2909
La Valdramina	12d03c03s01	2556
La Esperanza	12d03c01s01	2219
La Merced	12d03c04s01	2177
Control	12d03c02s02	2107
Galo Plaza	12d03c05s02	1846
Divino Niño	12d03c06s01	1162
7 de Octubre	12d03c07s03	966
Guayacán	12d03c08s01	952
Control	12d03c02s01	863
La Merced	12d03c04s02	446
Divino Niño	12d03c06s03	189

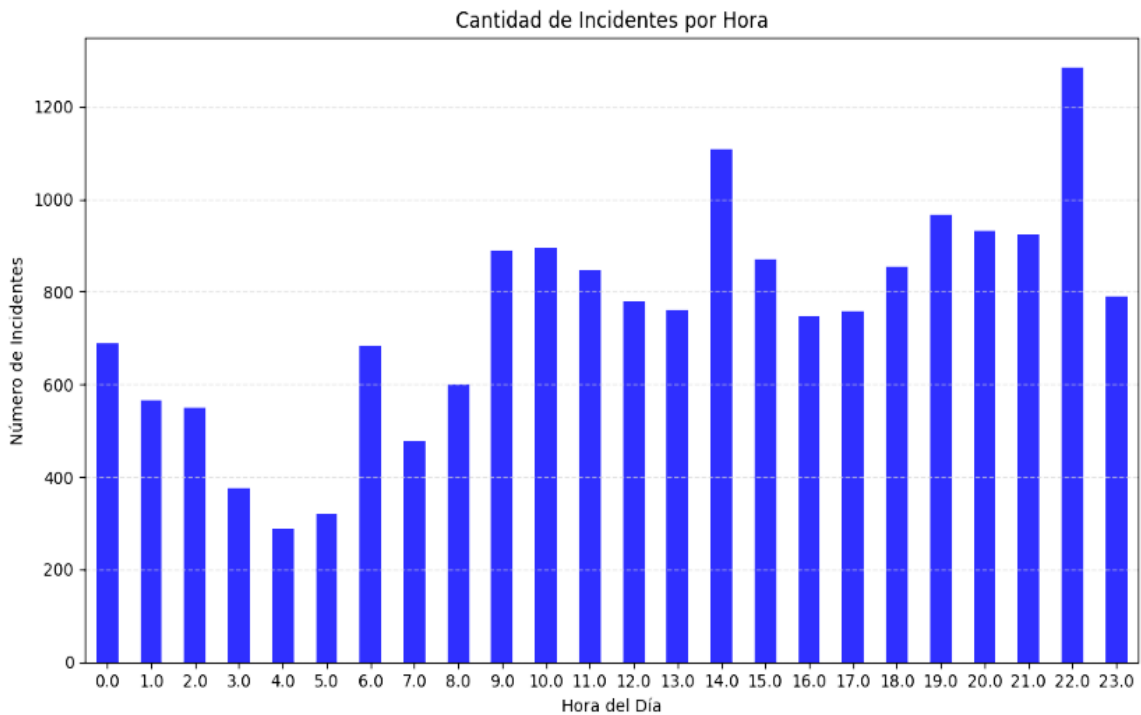
Se obtiene que los incidentes registrados por subcircuito reflejan una concentración predominante en las parroquias de 7 de Octubre, Divino niño 2 y Galo plaza

### 4.1.1.3. Horas Pico con Mayor Cantidad de Incidentes

Para el análisis de la data en Google Colab se extrajo la hora en la que ocurrió cada incidente, se los agrupó por el parámetro de hora, se los contó y ordenó de forma ascendente

**Figura 6.**

*Horas pico con mayores incidentes*



**Tabla 6.**

*Horas pico*

Hora	Número de incidentes
22.0	1284
14.0	1108
19.0	966
20.0	931
21.0	924

La tabla muestra las horas del día con mayor número de incidentes registrados, lo que permite identificar los momentos de mayor actividad en términos de emergencia.

#### 4.1.1.4. Incidentes más Frecuentes

**Tabla 7.**

*Incidentes más frecuentes*

Actividad	Número de incidentes
Patrullaje policial en el sector solicitado	13849
Resguardo a personas que traslada valores	3124
Presencia policial	2075
Libadores	1703
Ficha de datos información	1462

La tabla muestra que el patrullaje policial es la actividad con mayor incidencia de incidentes, reflejando su importancia central en la prevención y respuesta a emergencias. El tipo de incidente resguardo de valores posee 3124 incidentes, colocándose en segunda ubicación en el ranking de frecuencia, mostrando la importancia que la ciudadanía del sector le da a la protección del tránsito de valores en áreas de alto riesgo. La presencia policial, aunque con menor frecuencia, juega un papel crucial en la disuasión de delitos y la mejora de la seguridad pública.

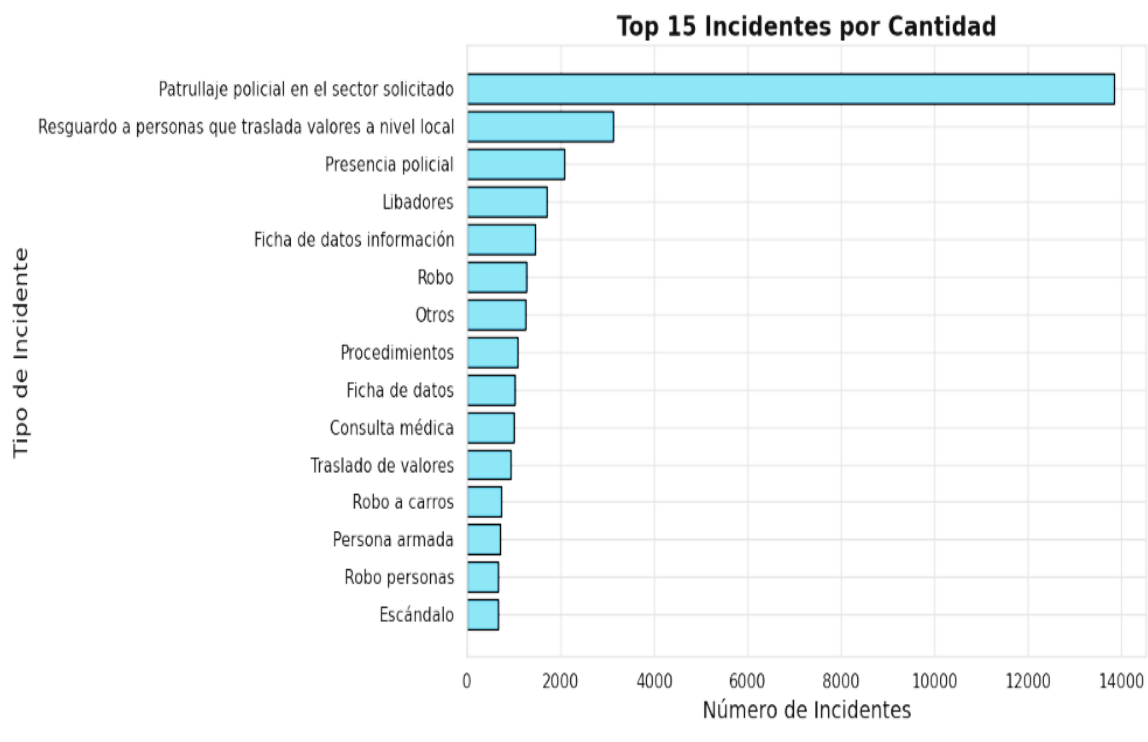
La presencia de libadores juega un papel representativo en la frecuencia de reportes de esta novedad, lo que orienta una cultura baja de respeto de los espacios públicos. Finalmente, la ficha de datos de información refleja un esfuerzo para recopilar información esencial, ayudando en la gestión y planificación de estrategias de seguridad.

De acuerdo con los datos facilitados por el sistema integrado de seguridad Ecu 911, se evidencia que los incidentes registrados en la ciudad son un total de 312 categorías, de los

que en resumen en la figura 7 se observa el top 15 de los incidentes de mayor frecuencia, siendo la categoría de patrullaje la de mayor demanda.

**Figura 7.**

*Top 15 de incidentes*



#### 4.1.1.5. Relación entre Circuitos e Incidentes

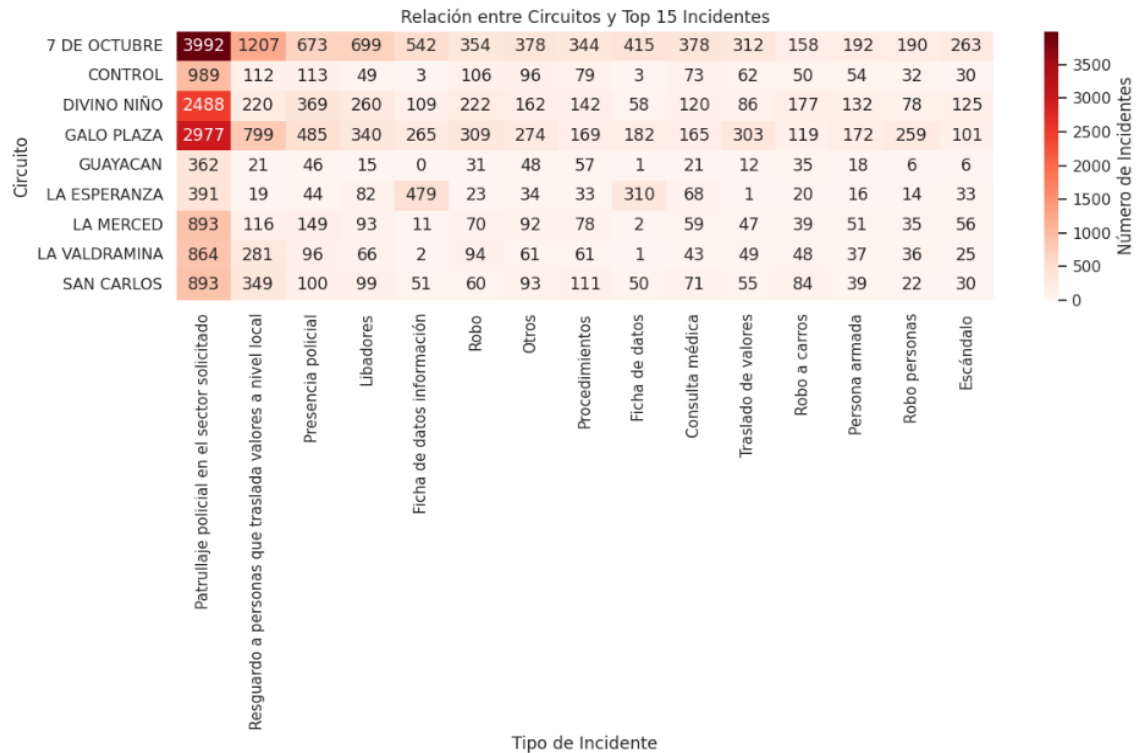
La relación entre los circuitos y los 15 tipos principales de incidentes en Quevedo destaca la parroquia 7 de Octubre, que es el circuito con la mayor cantidad de incidentes, especialmente en actividades como el patrullaje policial y resguardo de valores. Sin embargo, se observa que en la parroquia Galo Plaza y Divino Niño tienen una alta frecuencia de incidentes relacionados con presencia policial y libadores. Por otra parte, La Esperanza destaca con un número considerable de incidentes en robos.

En general, los incidentes más comunes están asociados a actividades de patrullaje y seguridad, mientras que los robos y otros tipos de incidentes también son frecuentes en varias

áreas. Este patrón resalta la importancia de asignar recursos de seguridad según las áreas con mayor incidencia de ciertos tipos de incidentes.

**Figura 8.**

*Mapa de calor de distribución de incidentes por circuito y tipo de incidente en Quevedo*



#### 4.1.2. Herramientas y Técnicas Utilizadas

Se empleó el lenguaje de programación Python mediante la plataforma Google Colab debido a su versatilidad, la amplia disponibilidad de bibliotecas enfocadas en el análisis de datos, y su facilidad para integrar herramientas de visualización y análisis estadístico. En este contexto, Python permitió la lectura del archivo Excel, el preprocesamiento de los datos, el cálculo de estadísticas descriptivas, la generación de gráficos y, posteriormente, la aplicación de algoritmos de clustering.

Google Colab permite la ejecución del software python directamente en la nube, sin la necesidad de tener el software instalado en la computadora. Además, ofrece un entorno amigable y compartido en el que es sencillo acceder a datos almacenados en Google Drive

y poder revisar o continuar los avances desde distintos computadores y lugares. Gracias a Colab se pudo:

- Realizar la lectura, carga y proceso del archivo de datos del ECU 911.
- Ejecutar el código para el análisis en un entorno interactivo.
- Aprovechar los recursos computacionales de manera gratuita, simplificando el proceso de ejecución y visualización de resultados.

Se utilizaron mapas de calor para visualizar y analizar la relación entre diferentes variables (por ejemplo, circuitos y tipos de incidentes) de una forma intuitiva. Esta técnica permite:

- Observar patrones de concentración de incidentes en ciertos circuitos.
- Facilitar la identificación de qué tipos de incidentes eran más frecuentes en determinados sectores.
- Comparar la intensidad (frecuencia) de cada categoría de incidente de una sola mirada.

Las técnicas de clustering, K-Means, DBSCAN o HDBSCAN son fundamentales.

Estos algoritmos ayudan a:

- Agrupar parroquias con características similares respecto a la frecuencia y tipo de incidentes.
- Detectar patrones escondidos que no se evidencian a simple vista.
- Determinar zonas vulnerables y de alta peligrosidad

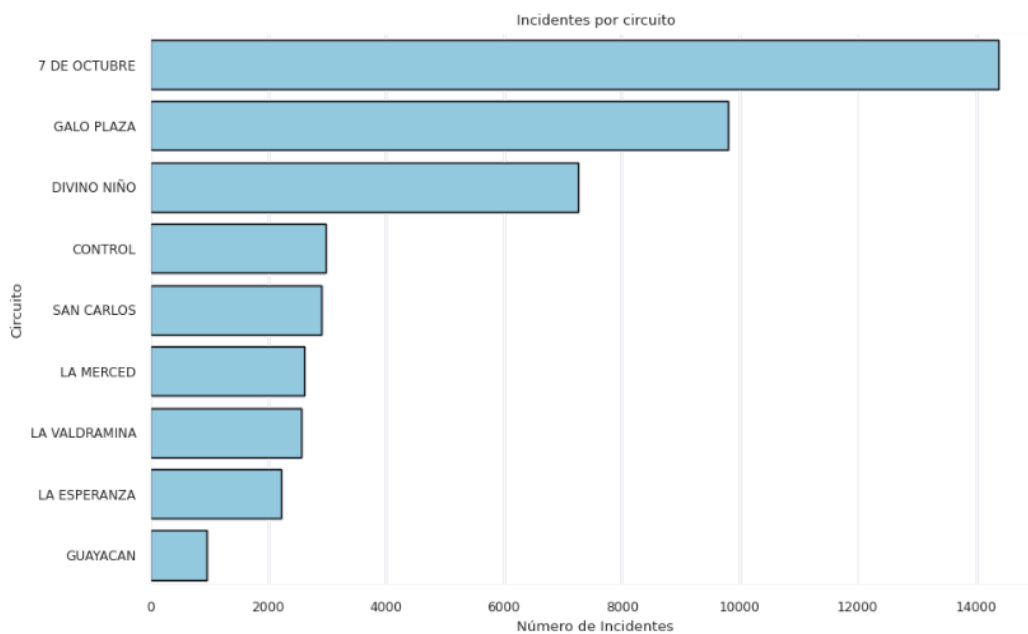
En conjunto, el uso de Python, Google Colab, y las distintas técnicas de visualización y el posterior uso de algoritmos de clustering brindan una metodología sólida para

transformar datos crudos en información valiosa. De esta forma, se obtienen resultados que orientan la toma de decisiones en seguridad pública y la asignación de recursos de manera más precisa.

### 4.1.3. Análisis Espacial

**Figura 9.**

*Incidentes por parroquia*



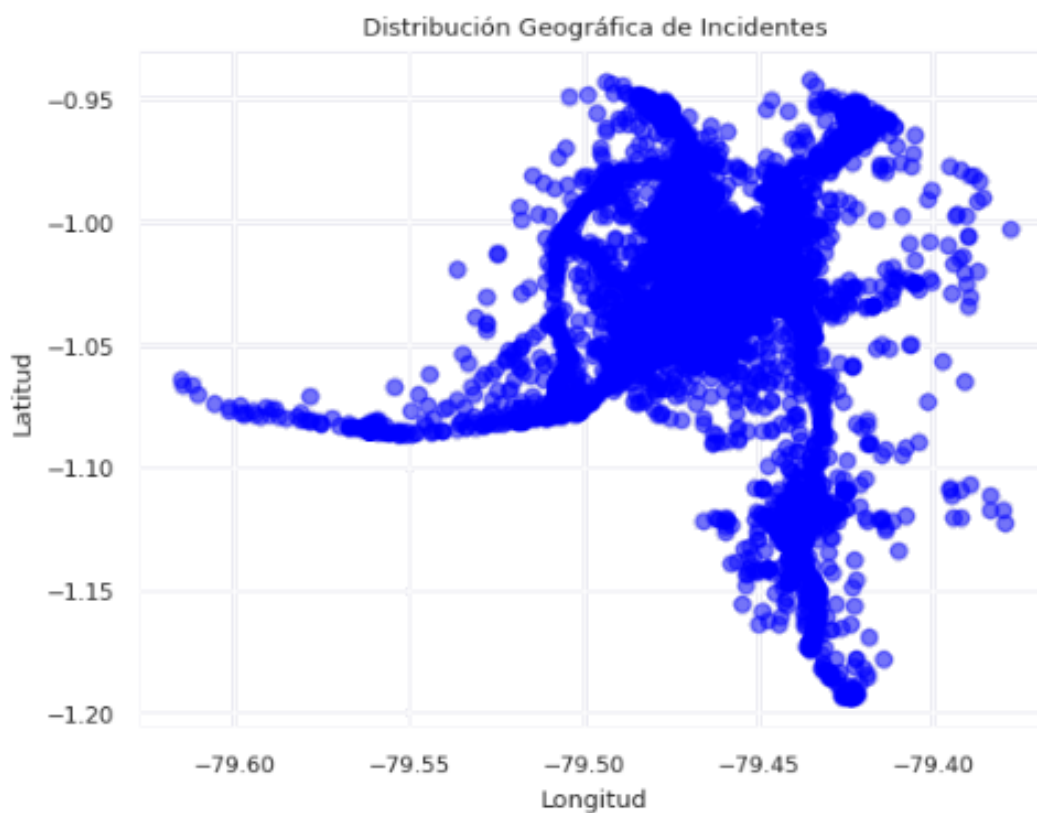
La gráfica muestra la distribución de los incidentes por circuito en Quevedo, destacando las zonas con mayor frecuencia de eventos. El circuito 7 de Octubre posee la mayor cantidad de incidentes reportados superando los 13.000, lo que refleja su alta densidad de población y actividad comercial. Le siguen el circuito de Galo Plaza y Divino Niño, con cifras que también indican una alta concentración de incidentes mayor a los 7000 y 9000 incidentes respectivamente, lo que sugiere que estas áreas son de alta movilidad y con mayor exposición a emergencias. Los circuitos como Control, San Carlos, La Merced, La Valdramina, La Esperanza y Guayacán presentan una frecuencia más baja de incidentes lo

que puede estar relacionado con una menor densidad de población o una menor actividad social.

Este análisis espacial resalta la necesidad de asignar recursos de manera estratégica y focalizada, priorizando los circuitos con mayor incidencia para optimizar la respuesta ante emergencias.

**Figura 10.**

*Distribución geográfica de los incidentes*



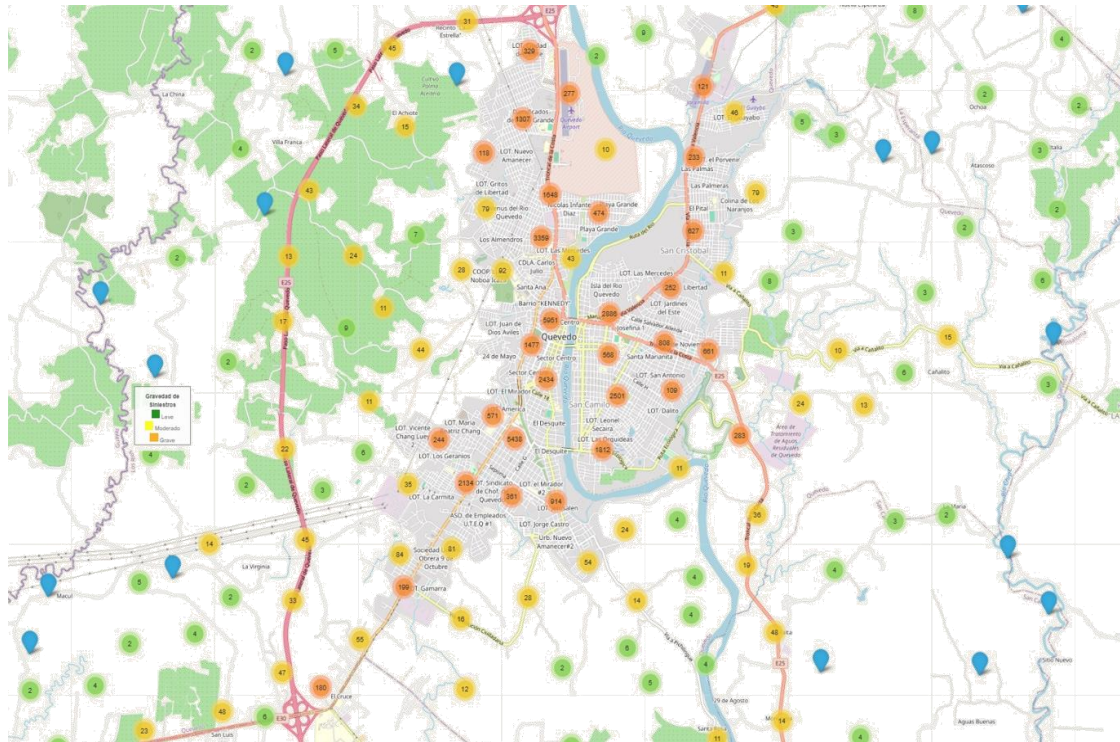
Los puntos azules dispersos a lo largo del gráfico indican la localización de los incidentes en la ciudad. Se puede observar una concentración significativa de puntos en ciertas áreas, lo que sugiere que los incidentes tienden a ocurrir principalmente en zonas urbanas o en áreas con mayor actividad, como en el centro de la ciudad y sus alrededores. La forma y densidad de la distribución refleja posibles hotspots o zonas de mayor riesgo, lo



que puede ser útil para dirigir esfuerzos de seguridad y recursos hacia esas áreas con mayor incidencia de eventos.

**Figura 11.**

*Distribución de puntos en la ciudad*



Los puntos en azul representan las áreas con menor gravedad, mientras que los puntos amarillos y naranjas muestran una referencia de mayor gravedad, permitiendo visualizar una mayor concentración o frecuencia de incidentes graves en esas áreas.

Se puede observar que las zonas con mayor gravedad están principalmente en el centro de Quevedo, y hacia las áreas periféricas se distribuyen la incidencia de casos es menor. En la gráfica, se visualiza que la zona urbana es la que mayormente concentra los casos reportados, por otra parte, la zona rural es en menos cantidad.

Los puntos verdes representan zonas con menor frecuencia de incidentes y se distribuyen en áreas más alejadas del centro de la ciudad.

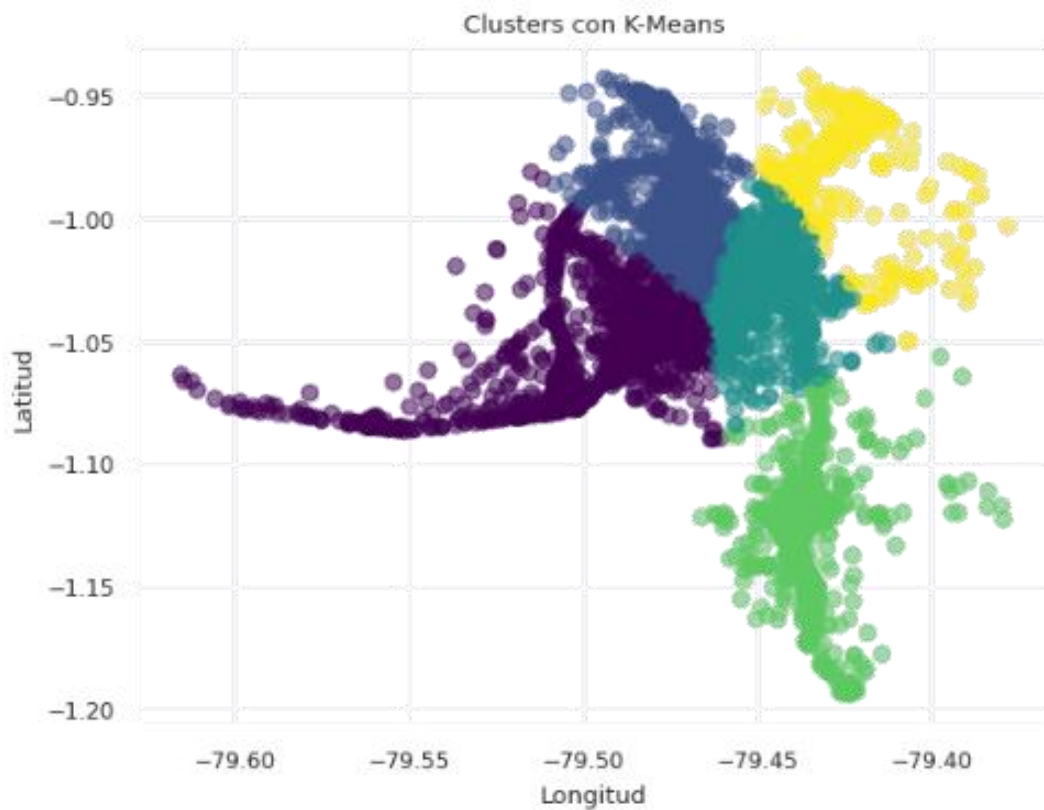
#### 4.1.3.1. Aplicación de Algoritmos de Clustering.

La finalidad de estos algoritmos es agrupar los datos con características similares, ya sea en términos espaciales (circuitos con perfiles parecidos a incidentes), temporales (horas o días con patrones comunes), o una combinación de ambas dimensiones.

#### KMEANS

Figura 12.

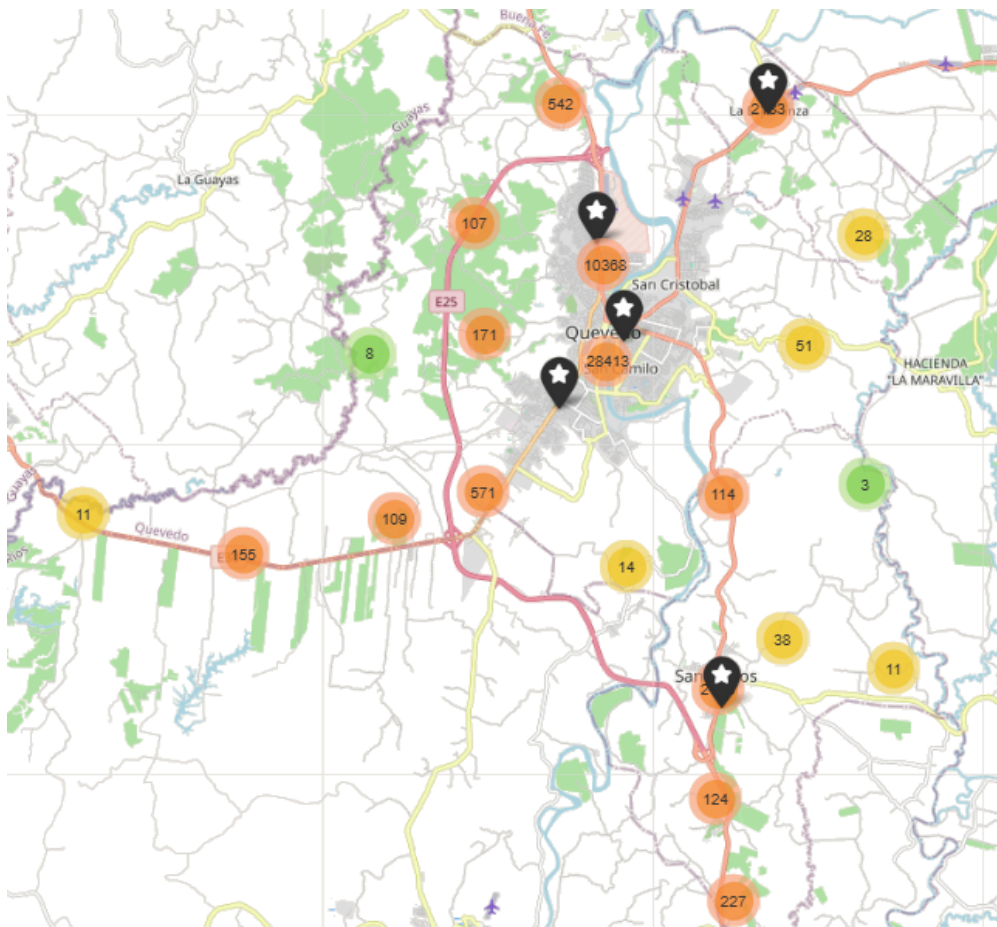
Gráfica algoritmo Kmeans



La elección de 5 clúster fue determinada para equilibrar la granularidad de la segmentación y la claridad en la interpretación de los resultados. Un número menor de clústeres podría haber simplificado demasiado la representación, mientras que un número mayor podría haber fragmentado excesivamente las áreas, dificultando la interpretación. El número 5, en este caso, permite una segmentación efectiva que resalta tanto las zonas de alta concentración como las áreas más dispersas de incidentes.

**Figura 13.**

*Centroides Kmeans*



Los centroides de cada cluster se añaden al mapa, representando el centro de cada agrupación de incidentes. Estos centroides son puntos clave para identificar áreas de alta gravedad o concentración de eventos. Al ubicarlos en el mapa con la ayuda de folium, se facilita la visualización espacial del clúster, proporcionando información valiosa sobre la distribución geográfica de los incidentes y permitiendo la toma de decisiones informadas sobre la asignación de recursos y la priorización de áreas críticas.

### **DBSCAN**

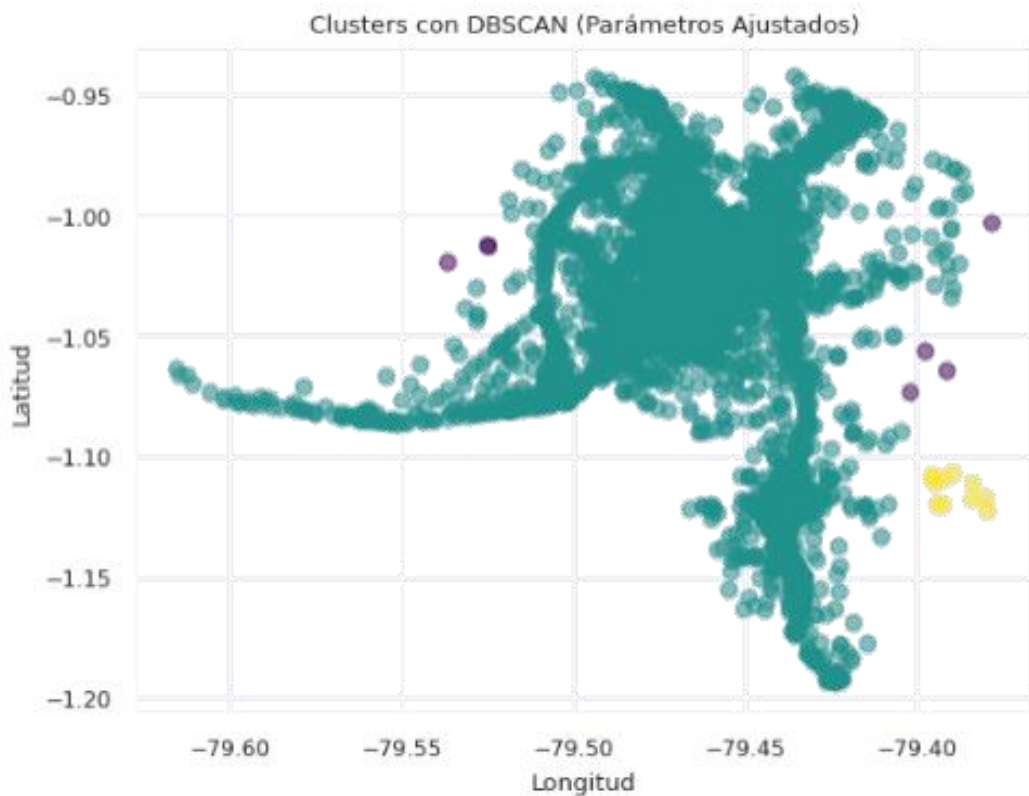
El algoritmo DBSCAN no solo permite identificar áreas con alta densidad de incidentes, sino que también es eficaz para encontrar patrones geográficos ocultos en los

datos. Al no requerir que se especifique un número fijo de clúster, el algoritmo DBSCAN realiza su agrupación de los datos según la distribución real de los incidentes.

Al identificar las áreas con actividad irregular, se obtiene una visión más clara de las zonas donde no se observan patrones evidentes, lo que facilita la optimización de recursos y la toma de decisiones más acertadas en las intervenciones de seguridad. Este enfoque flexible y adaptable ofrece una herramienta eficaz para gestionar incidentes en tiempo real, permitiendo a los encargados de la seguridad actuar de manera más informada y eficiente

**Figura 14.**

*Algoritmo DBSCAN*



En la gráfica de salida se visualiza que la mayoría de los incidentes pertenecen al clúster 0 (45,656 incidentes), lo que indica que de acuerdo al análisis del algoritmo los incidentes se encuentran concentrados en un único grupo.

El parámetro  $\text{eps}=0.1$  utilizado en DBSCAN define la distancia máxima para considerar que dos puntos están en el mismo clúster, mientras que  $\text{min\_samples}=5$  establece el número mínimo de puntos requeridos para formar un clúster denso.

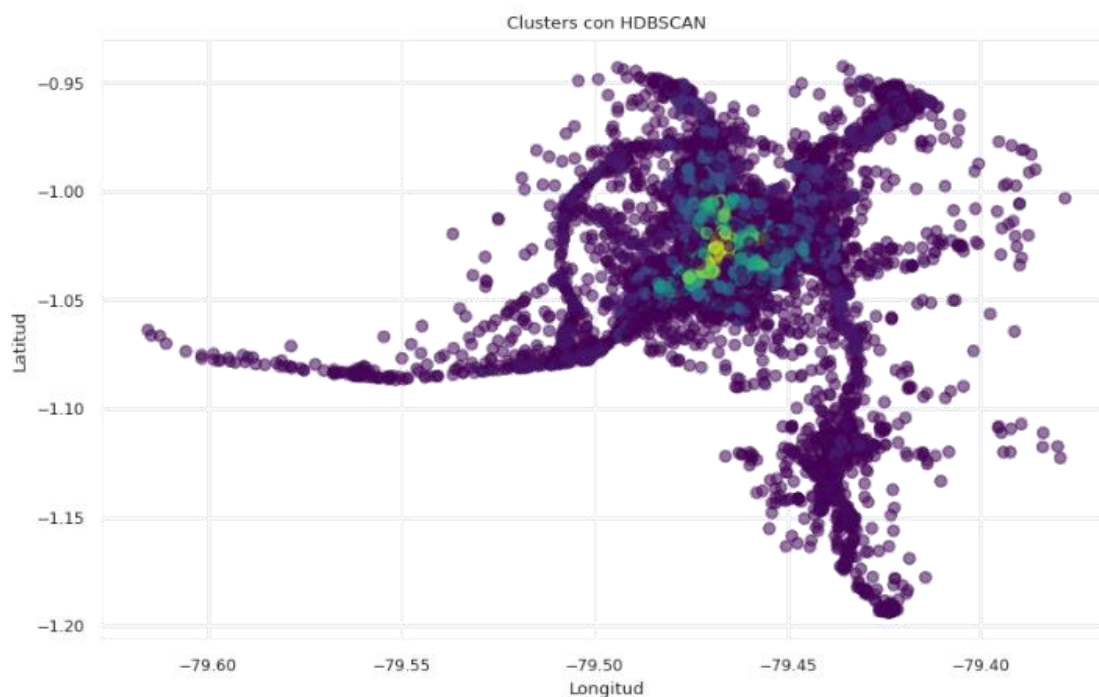
Los puntos amarillos corresponden a los puntos etiquetados como -1, que son los puntos de ruido. Estos puntos no forman parte de un clúster denso.

## HDBSCAN

Permite agrupar los incidentes en función de sus coordenadas geográficas utilizando de la data las columnas de latitud y longitud de cada incidente reportado. Los puntos agrupados en clúster se visualizan en un gráfico de dispersión, donde los colores indican a qué clúster pertenece cada punto. Los puntos que no pertenecen a ningún clúster denso son etiquetados como -1 y se muestran en la gráfica de forma diferenciada, lo que permite visualizar zonas de alta y baja concentración de incidentes.

**Figura 15.**

*Algoritmo HDBSCAN*



Este código aplica HDBSCAN para agrupar los incidentes en función de sus coordenadas geográficas (latitud y longitud). Los puntos agrupados en clúster se visualizan en un gráfico de dispersión, donde los colores indican a qué cluster pertenece cada punto. Los puntos que no pertenecen a ningún cluster denso se etiquetan como -1 y se muestran de forma diferenciada, lo que permite visualizar zonas de alta y baja concentración de incidentes

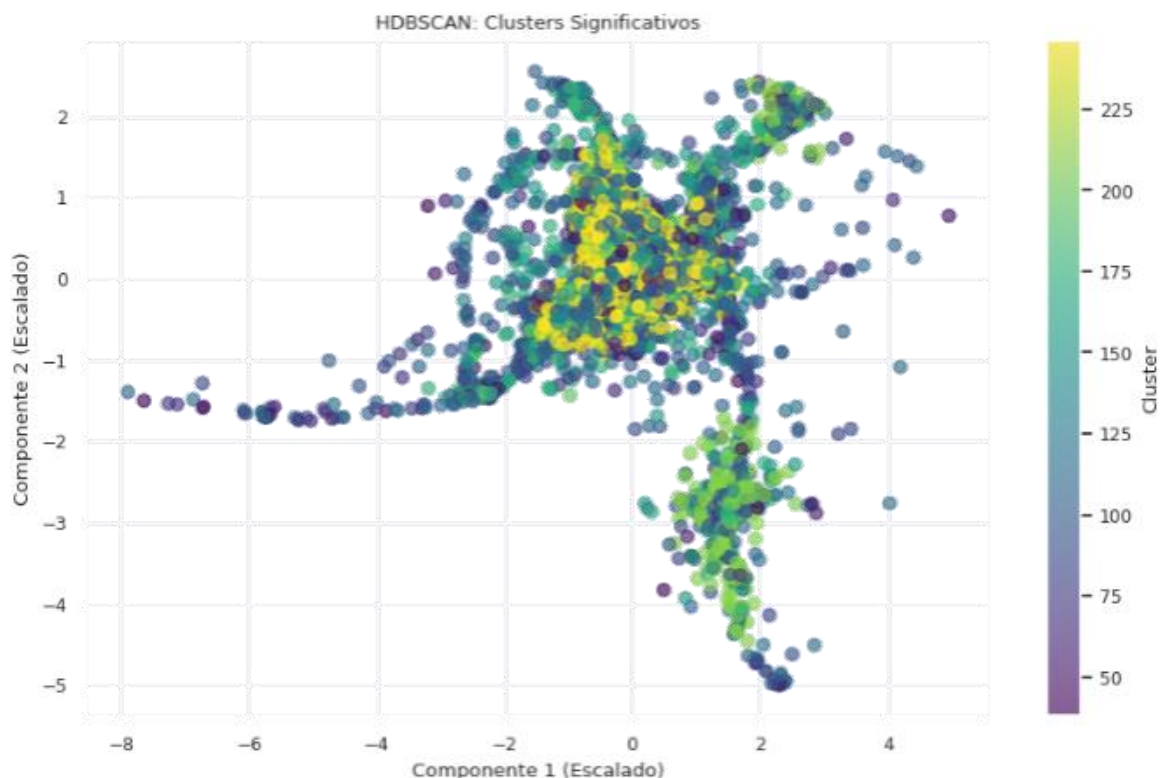
Mediante código:

```
# Filtrar clúster significativos
significant_clúster = {label: size for label, size in
cluster_sizes.items() if size > 50 and label != -1}
print(f"Clúster significativos: {significant_clúster}")
```

Permite tener una mejor visualización de los clústeres solo graficando los puntos de más de 50 incidentes agrupados.

**Figura 16.**

*HDBSCAN clústers significativos*



#### 4.1.3.2. Evaluación de Algoritmos.

Se evalúan tres algoritmos K-Means, DBSCAN y HDBSCAN, utilizando dos métricas comunes para evaluar la calidad del clustering: Silhouette Score y Davies-Bouldin Index (DBI). El objetivo es comparar el desempeño de los tres algoritmos en cuanto a la calidad de los clústeres generados. Se utilizan dos métricas:

**Silhouette Score:** Mide la cohesión de los puntos dentro de su propio clúster y la separación con otros clústeres. Los valores van de -1 a 1, donde 1 significa que los puntos están bien agrupados y alejados de otros clústeres, y -1 significa que los puntos están mal agrupados. Mayor es mejor.

**Davies-Bouldin Index (DBI):** Mide la dispersión de los clústeres. Un valor bajo de DBI indica una buena separación entre clúster. Menor es mejor.

Pasos del Código:

1. Carga y Preprocesamiento de Datos:
2. Se asegura de que los datos no contengan valores NaN con `dropna()`.
3. Se normalizan los datos utilizando `StandardScaler` para que todas las características tengan la misma escala antes de aplicar los algoritmos de clustering.

Clustering con K-Means:

4. K-Means se aplica con 5 clúster (`n_clúster=5`).
5. Después de ajustar el modelo, se calcula el Silhouette Score y DBI si se detectan al menos 2 clúster distintos, es decir, si no hay solo un único cluster.
6. Los resultados se almacenan en un diccionario resultados.

Clustering con DBSCAN:

7. DBSCAN se ajusta utilizando parámetros  $\text{eps}=0.5$  y  $\text{min\_samples}=5$ , que son configuraciones estándar para comenzar.
8. Al igual que en K-Means, se evalúa el Silhouette Score y DBI solo si hay más de un clúster distinto y no todos los puntos son ruido (-1).

Clustering con HDBSCAN:

9. HDBSCAN se aplica con parámetros  $\text{min\_cluster\_size}=5$  y  $\text{min\_samples}=5$ .
10. Se calcula el Silhouette Score y DBI solo para las muestras que no son ruido (las que no tienen etiqueta -1).

Resultados obtenidos de la métrica:

**Tabla 8.**

*Resultados obtenidos de la métrica*

	Silhouette	DBI
K-Means	-0,06	0,74
DBSCAN	nan	nan
HDBSCAN	0,52	0,69

Los resultados obtenidos de la métrica Silhouette Score y DBI arrojan valores claros que permiten realizar una selección adecuada para determinar el mejor algoritmo a utilizar en el análisis de los datos entregados por el servicio integrado de seguridad ECU 911 y que agrupa los incidentes suscitados durante al año 2023 en la ciudad de Quevedo.

Comparativo e Interpretación de los Resultados

De acuerdo con los resultados obtenidos en la tabla 8 se observa que:



**Tabla 9.**

Comparativo e Interpretación de los Resultados

	Silhouette	DBI
K-Means	Un valor negativo de Silhouette sugiere que los puntos dentro de los clústeres no están bien agrupados y que podrían estar más cerca de otro clúster. Esto indica que K-Means no ha realizado un buen trabajo de agrupamiento en este caso específico, ya que los clústeres no son lo suficientemente coherentes ni separados.	Un valor de DBI bajo indica que los clústeres generados por el algoritmo K-Means tienen una separación aceptable, pero la calidad del clustering es inferior en comparación con los otros algoritmos.
DBSCAN	El valor NaN sugiere que DBSCAN ha generado solo un clúster o ha considerado todos los puntos como ruido (etiquetados como -1). Como resultado, no se puede calcular un Silhouette útil en este caso	Dado que DBSCAN no generó múltiples clústeres, es decir generó ruido en su lugar, no es posible calcular un DBI válido.
HDBSCAN	Un valor positivo y moderado indica que HDBSCAN ha creado clúster bien definidos, con una buena cohesión interna y una separación adecuada entre clúster. En cuanto más cercano a 1 esté, mejor es la calidad del clustering, por lo que HDBSCAN está proporcionando una segmentación más adecuada que K-Means en este caso.	El valor de DBI es bajo, lo que indica que los clústeres generados por HDBSCAN tienen una buena separación. Al ser este menor a 1 significa que los clústeres están bien definidos y están bien separados.

#### **4.1.4. Análisis Temporal**

En el contexto de esta investigación, el análisis temporal tiene como objetivo identificar los patrones de ocurrencia de incidentes en el servicio integrado de seguridad ECU 911 en Quevedo a lo largo del tiempo. A través de este análisis, se examinan las incidencias según la hora del día, el día de la semana y el mes, permitiendo detectar las horas y días con mayor frecuencia de eventos. Como también el evolutivo a través del tiempo, permitiendo evidenciar los meses más violentos. Este conocimiento es crucial para identificar los períodos de mayor actividad y aquellos con menor incidencia, relacionándolos con sectores geográficamente divididos en parroquias.

Al entender estos patrones, se facilita la asignación eficiente de recursos y la planificación de intervenciones de seguridad durante los momentos de mayor demanda, optimizando así la respuesta ante emergencias y mejorando la efectividad de las estrategias de seguridad en la ciudad.

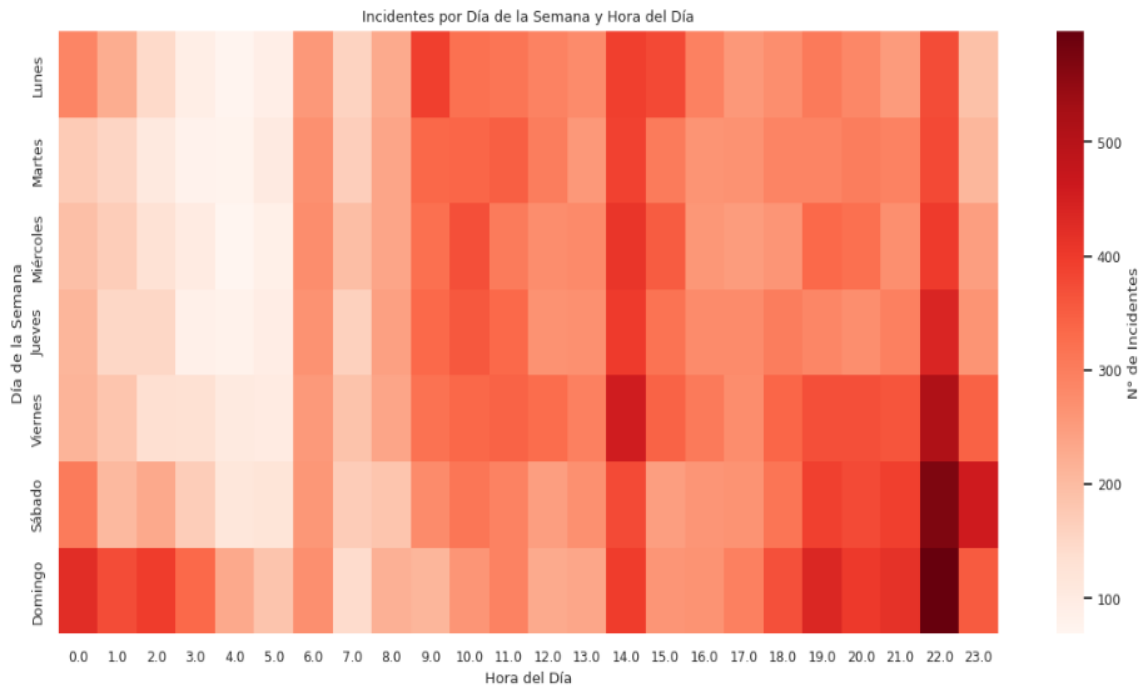
##### **4.1.4.1. Mapa de Calor de los Incidentes por Día de la Semana y Hora del Día.**

El presente mapa de calor muestra la distribución de los incidentes registrados según el día de la semana y la hora del día. Esta visualización permite identificar patrones temporales de los incidentes, destacando los días y horas con mayor frecuencia de eventos. Este análisis ayuda a comprender las tendencias temporales y puede ser útil para asignar recursos de manera más eficiente durante los períodos de mayor demanda.

Además, de realizar la asignación de personal en los sitios de mayor necesidad y en las horas adecuadas que permitan dar una respuesta oportuna a las necesidades de la ciudadanía, permitiendo disminuir los índices de inseguridad y optimizar de forma eficiente y efectiva los recursos asignados a los departamentos involucrados en las respuestas de incidentes en la ciudad.

**Figura 17.**

*Incidentes por día de la semana y hora del día*



Horas con mayor incidencia de incidentes:

Hora	Count
22	3277
14	2825
19	2420
20	2332
10	2302

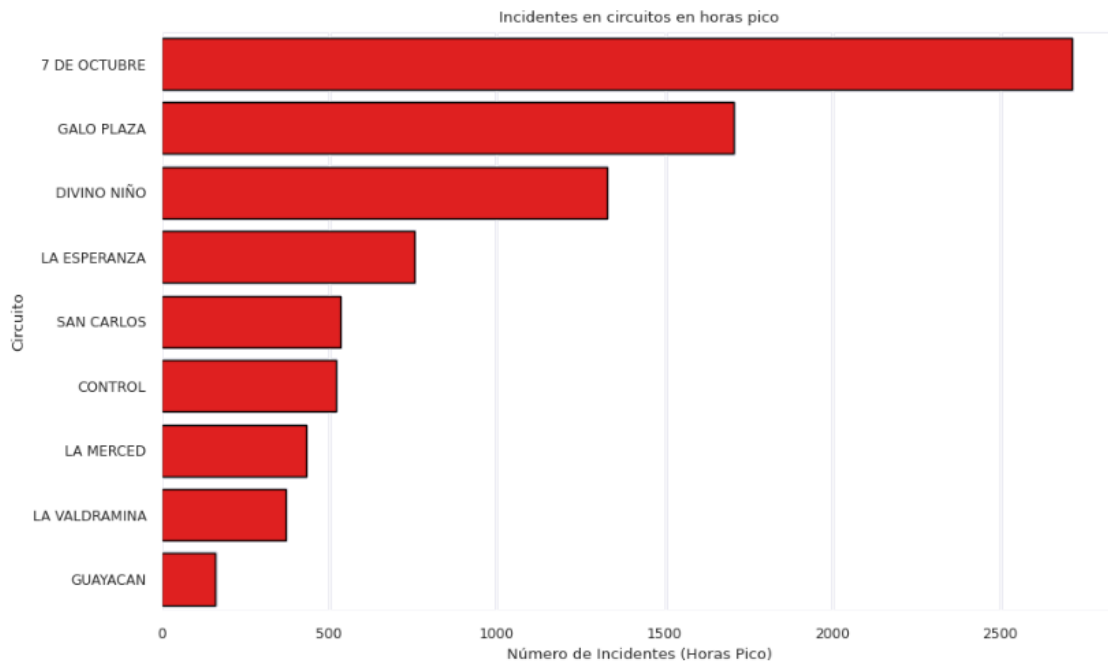
En la figura 18 se evidencia que las 5 horas de mayor demanda de incidentes en las distintas parroquias son las 10 am, 14 pm, 19 pm, 20 pm y 22 pm, lo que se considera que son los momentos de mayor índice delictual de acuerdo a los reportes realizados en el ECU 911.

#### 4.1.5. Incidentes en Circuitos en Horas Pico

Considerando las horas determinadas en el enunciado anterior como horas pico donde se dan la mayor cantidad de eventos se evalúa con los circuitos en la que se da la mayor cantidad de eventos, teniendo como resultado:

**Figura 18.**

*Incidentes en circuitos en horas pico*



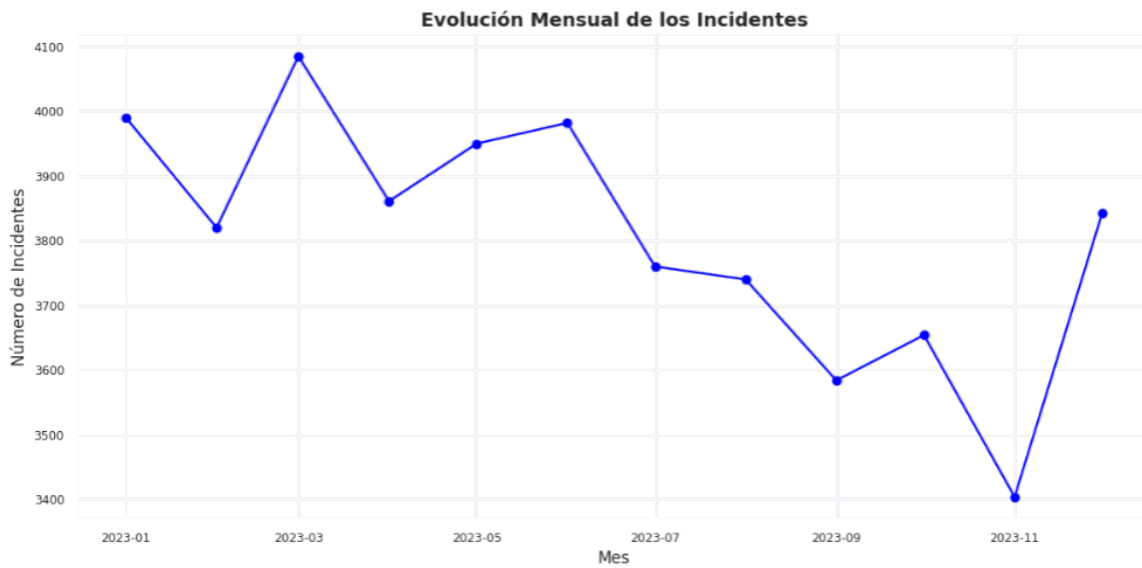
En la figura 19, se evidencia 3 circuitos en los cuales se desarrollan durante la hora pico la mayor cantidad de eventos delincuenciales, teniendo a la parroquia 7 de octubre encabezando la lista.

#### **4.1.6. Evolución Mensual**

La evolución mensual de los incidentes permite evidenciar y tener una visión general de cómo varía la frecuencia de eventos a lo largo del tiempo. Este análisis identifica las tendencias y patrones estacionales en el desarrollo de los incidentes, mostrando si los meses con mayor y los de menor cantidad de eventos. Al observar la variación de eventos en cada mes, se pueden detectar picos de actividad, lo cual puede estar relacionado con factores específicos como el clima, festividades, o cambios en la seguridad y movilidad de la población. Esta información es clave para asignación de recursos y la toma de decisiones de forma estratégicas ante la gestión de emergencias, asegurando que se pueda responder de manera efectiva durante los períodos de mayor demanda.

**Figura 19.**

*Evolución Mensual de los incidentes*

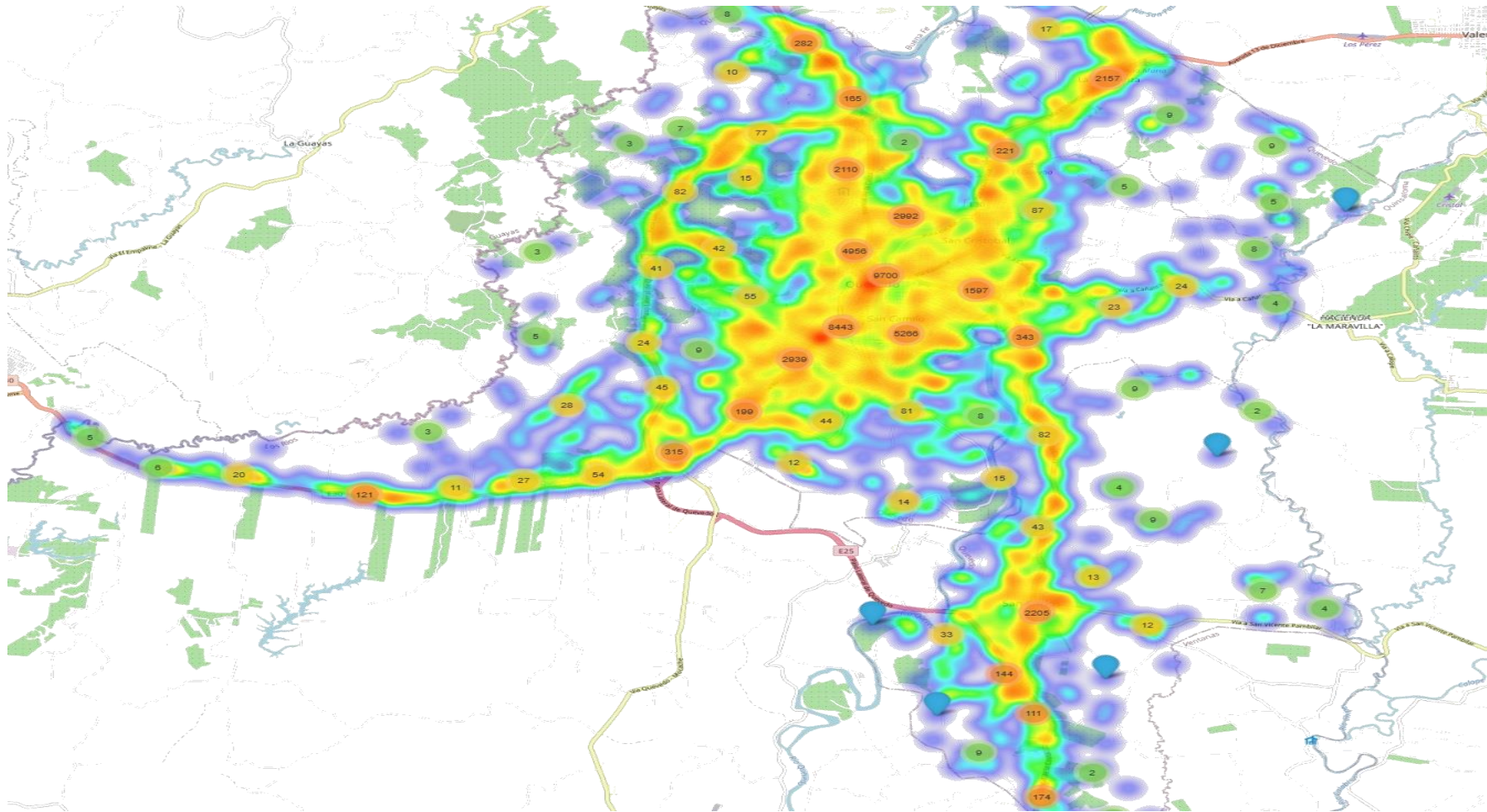


Se evidencia fluctuaciones en el número de incidentes registrados a lo largo del año 2023. Se observa un aumento significativo en los meses de enero, marzo y noviembre, con picos por encima de los 4,000 incidentes. En los meses de junio a septiembre del 2023 existe una disminución notable de incidentes bajando a menos de 3,800. Este comportamiento sugiere que los incidentes tienen una variabilidad estacional, con ciertos meses presentando más actividad, lo que podría estar influenciado por factores externos como el clima, eventos especiales o cambios en la seguridad pública, no obstante, se observa un pico considerable hacia el mes de diciembre, que podría estar relacionado a los eventos que se desarrollan en este mes, teniendo un cierre de año con menor índice delincencial que el inicio.

#### 4.1.7. Mapa de Calor

Figura 20.

Mapa de calor puntos calientes



## **Análisis:**

- Zonas de alta densidad (colores cálidos): Se observan áreas con una alta concentración de incidentes orientadas al centro de la ciudad de Quevedo y las zonas cercanas a las principales vías de acceso. Las áreas de color amarillo, rojo y naranja indican los lugares con el mayor número de incidentes durante el año, lo que refleja una mayor actividad en esas zonas, debido a la densidad poblacional y flujo vehicular.
- Zonas de baja densidad (colores fríos): Las áreas con colores azules y verdes muestran una baja concentración de incidentes en comparación a la zona centro, teniendo que las periferias de la ciudad tienden a presentar menos frecuencia de incidentes de emergencias.
- Puntos de alta ocurrencia (números en los círculos): En el mapa, los números dentro de los círculos indican el número específico de incidentes que se han dado en los lugares aledaños y son puntos clave para el análisis delincencial dentro del sector. Estos números permiten identificar con precisión las ubicaciones de mayor frecuencia de incidentes.
- Posibles patrones espaciales: La gráfica también permite detectar posibles patrones espaciales, como la relación entre la ubicación donde se dan los incidentes y las vías principales de transporte, lo que es útil para asignar recursos y mejorar las estrategias de respuesta a emergencias.
- Áreas críticas: Las zonas con la mayor concentración de eventos, especialmente en las áreas cercanas a las principales carreteras y el centro urbano, son áreas críticas que requieren un mayor enfoque en términos de seguridad y gestión de emergencias.

## **4.2. Discusión de los Resultados**

En los apartados anteriores se presentó el análisis de los patrones temporales, espaciales y de la evolución de los incidentes en Quevedo. A partir de los resultados obtenidos, se destacan varias conclusiones relevantes que permiten comprender mejor la dinámica de los incidentes y las áreas que requieren atención prioritaria.

En primer lugar, se observará una variabilidad considerable en el número de incidentes registrados durante el año 2023. Teniendo un aumento notable en los meses de enero, marzo y noviembre, con picos que superaron los 4,000 eventos. Estos picos pueden estar vinculados a factores externos, como el clima, festividades o un incremento en la actividad social, que son factores que generan un mayor número de emergencias. En los meses de junio, julio y septiembre los incidentes bajaron considerablemente, lo que indica que durante este período existió una baja actividad delincuencia o una mejora en la respuesta ante emergencias.

El análisis espacial realizado en los incidentes delictivos permitió identificar patrones claros en la distribución de los delitos dentro de la ciudad de Quevedo. A través del mapa de calor, se evidencia una alta concentración de incidentes en el centro de la ciudad y a lo largo de las principales vías de transporte. En la zona urbana, las parroquias 7 de Octubre y Galo Plaza presentan la mayor cantidad de índices delictivos, seguidas por Divino Niño, mientras que, en la zona rural, la parroquia La Esperanza destaca como un punto crítico de inseguridad, especialmente en la categoría de robos.

Estos hotspots delictivos reflejan la relación entre la criminalidad y factores como la alta densidad poblacional, la cercanía a las rutas de acceso de la ciudad y el flujo constante de personas y vehículos, así como la actividad comercial de estas zonas. El análisis respalda la necesidad de una asignación estratégica de recursos en estos puntos críticos, priorizando patrullajes, vigilancia y medidas preventivas adaptadas a las dinámicas de cada sector.



Además, los hallazgos proporcionan evidencia concisa para la formulación de políticas de seguridad ciudadana, permitiendo a las autoridades optimizar sus estrategias de intervención con base en datos reales y actualizadas. La visualización geoespacial de los incidentes ocurridos facilita la identificación de tendencias y la toma de decisiones de los entes gubernamentales para reducir los tiempos de respuesta ante emergencias y mitigar el impacto de la delincuencia en las áreas más afectadas.

Adicionalmente, al comparar los algoritmos de clustering aplicados (K-Means, DBSCAN y HDBSCAN), los resultados mostraron que HDBSCAN fue el mejor algoritmo en términos de Silhouette Score y Davies-Bouldin Index (DBI), lo que muestra que este algoritmo es más efectivo para detectar clúster con mayor separación y cohesión en los datos espaciales de los incidentes. K-Means, en cambio, mostró un rendimiento inferior, con un Silhouette Score negativo, lo que muestra que los clústeres generados no son tan coherentes ni están bien separados.

Es importante destacar que los patrones espaciales y temporales de los incidentes permiten optimizar las estrategias de seguridad y asignación de recursos de los diferentes entes de control. Las zonas con alta densidad de incidentes deben recibir atención prioritaria, y la programación de patrullajes o intervenciones deben ser ajustada según los picos de actividad detectados en los meses de mayor frecuencia de eventos.

El análisis temporal de los incidentes delictivos revela patrones significativos en la distribución horaria de los eventos. Los resultados muestran que las horas con mayor concentración de incidentes son las 22:00, 14:00 y 19:00, reflejando picos de actividad delictiva en la noche y en momentos clave del día.

Los actos delictivos no ocurren de manera aleatoria, estos están influenciados por demás factores como el aumento del tráfico vehicular, la concentración de personas en espacios públicos y la disminución de vigilancia en ciertas franjas horarias. La alta

incidencia en la noche puede estar relacionada con la menor presencia de transeúntes y patrullaje, mientras que los picos en la tarde coinciden con horarios de salida laboral y escolar, donde el movimiento en la ciudad es intenso.

Estos hallazgos son fundamentales para la realización de la planificación de estrategias de seguridad, ya que permiten optimizar la distribución de recursos y personal en los momentos críticos. Además, proporcionaría información clave para el diseño de políticas de prevención del delito, facilitando la implementación de operativos focalizados y proactivos que ayudan a reducir la criminalidad en los horarios más vulnerables.

En cuanto a los tipos de incidentes más comunes, los resultados muestran que el patrullaje policial (con 13,849 incidentes), el resguardo de personas que trasladan valores (3,124 incidentes), y la presencia policial (2,075 incidentes) son los incidentes más frecuentes dentro de los registros analizados. Esto refleja las principales acciones a implementarse relacionadas con la seguridad ciudadana y la protección de bienes y personas, indicando que las autoridades priorizan estas actividades en las áreas con mayor concentración de incidentes. Este análisis también subraya la importancia de continuar con estas estrategias, dado su impacto en la prevención y gestión de emergencias.

En resumen, los resultados de este análisis sugieren que la distribución de los incidentes en Quevedo sigue patrones tanto temporales como espaciales definidos. El uso de herramientas como el análisis de datos mediante clustering y la visualización con mapas de calor ofrece una poderosa herramienta para tomar decisiones en la gestión de emergencias, lo que permite tomar medidas para mejorar la eficiencia y efectividad de las respuestas del ECU 911 y entes asociados.

## Capítulo 5

### Marco propositivo

#### 5.1. Planificación de la Actividad Preventiva

En este contexto, se visualiza una oportunidad para mejorar la gestión de incidentes en la ciudad de Quevedo, utilizando herramientas tecnológicas avanzadas que permitan un análisis más efectivo de los patrones temporales y espaciales. La propuesta sugiere implementar modelos de clustering avanzados como HDBSCAN, los cuales ofrecen ventajas sobre otros métodos tradicionales al identificar patrones ocultos y zonas críticas, mejorando la toma de decisiones en la gestión de emergencias. Para el próximo periodo de operación del Sistema de seguridad integrado ECU 911 se sugiere considerar los resultados obtenidos con el objetivo de que mejore sus tiempos de respuesta y distribuyan de manera eficiente sus recursos, optimizando personal, equipos, vehículos, entre otros.

Además, se plantea la utilización de software de análisis geoespacial y herramientas de visualización como QGIS o ArcGIS, que permiten mapear los incidentes y realizar análisis predictivos en tiempo real. Estos softwares tienen la facilidad de realizar ajustes y mejoras continuas en la gestión de emergencias, integrando datos de diferentes fuentes y facilitando la toma de decisiones en situaciones críticas.

Para análisis más complejos, como la predicción de patrones de incidentes, se recomienda el uso de herramientas estadísticas y de programación como Python o R, que son utilizadas en estudios avanzados y por expertos en seguridad pública y gestión de emergencias. Estos lenguajes permiten realizar simulaciones y modelar escenarios futuros, lo que ayuda a una planificación más eficiente de los recursos, además, de una predicción de posibles eventos a suscitarse considerando no solo la ubicación geográfica sino también el momento en el que posiblemente puede ocurrir.

Con el fin de asegurar una implementación efectiva de estas tecnologías, es necesario llevar a cabo un programa de capacitación para los operadores y analistas de ECU 911, con un enfoque en la formación de uso de estas herramientas avanzadas y en la interpretación de los resultados obtenidos. Esto no solo contribuiría al desarrollo profesional de los trabajadores, sino que también mejora la eficiencia de las respuestas que se puedan dar ante emergencias, con la posibilidad de ajustar las estrategias a medida en al que estas suceden, ya que día a día la base datos se alimentara de nuevos incidentes y se identifican nuevas tendencias.

Este enfoque integral permitirá un desarrollo tanto en la capacidad operativa del personal como en la mejora de la seguridad y la eficiencia en la respuesta ante incidentes, alineándose con las mejores prácticas internacionales y los estándares de gestión pública de emergencias.

## Conclusiones

Mediante el análisis de los datos del servicio integrado de seguridad ECU 911 se identificaron patrones espaciales y temporales claros en la distribución de los incidentes. El uso de algoritmos de clustering con HDBSCAN permiten agrupar los incidentes en zonas específicas y determinar las parroquias con mayor índice delincencial.

El análisis de la estructura y calidad de los datos proporcionados por el ECU 911 Quevedo están bien organizados y son de alta calidad, lo que facilita su procesamiento. La información es clara, completa y coherente, lo que permite identificar fácilmente los elementos necesarios para su preprocesamiento. Con esta base sólida, se pueden realizar análisis más profundos y efectivos para mejorar la gestión de emergencias en la ciudad.

Se concluyó que los datos facilitados por el ECU 911 contenían registros valiosos y mediante la limpieza, normalización y la aplicación de técnicas de preprocesamiento como la eliminación de valores nulos, la estandarización del formato de las fechas mejoró la calidad de los datos para su análisis posterior.

El uso del algoritmo HDBSCAN facilitó una segmentación detallada de las zonas con alta concentración de incidentes, proporcionando información clave para la toma de decisiones en materia de seguridad y permitió identificar con precisión que la zona centro de la ciudad es la de mayor riesgo en Quevedo.

Para evaluar la calidad de los algoritmos, se emplearon métricas de desempeño como el Silhouette Score y el Índice de Davies-Bouldin (DBI). Obtenidos estos resultados se concluye que HDBSCAN en este estudio es el algoritmo más eficiente para identificar clústeres bien definidos, logrando una adecuada separación y cohesión entre los datos. Esto refuerza su utilidad para el análisis de incidentes y la toma de decisiones estratégicas en temas de seguridad ciudadana

## **Recomendaciones**

Se recomienda el uso de algoritmos de clustering para el análisis de cantidades considerables de datos donde se requiera agruparlos. Es importante realizar pruebas adicionales con diferentes parámetros de este algoritmo para ajustarlo aún más a las características específicas de los datos de incidentes.

Con el objetivo de aprovechar al máximo la calidad y organización de los datos del ECU 911, se recomienda implementar procesos automatizados de preprocesamiento y análisis mediante herramientas avanzadas como Python o R. Esto permitirá agilizar la extracción de información relevante, detectar tendencias de manera más eficiente y optimizar la toma de decisiones en la gestión de emergencias.

Se recomienda para la ampliación del estudio, y la aplicación de otros tipos de algoritmos se realice una recolección de datos más detallados, como la inclusión de factores socioeconómicos o eventos especiales, los cuales enriquecerían el análisis investigativo.

Se recomienda aplicar medidas de seguridad focalizada en la zona centro de la ciudad, donde se acumulan la mayor cantidad de incidentes, redistribuir los recursos de acuerdo a las necesidades y las franjas horarias de altos índices delincuenciales.

Dado que HDBSCAN dentro de los 3 algoritmos estudiados ha demostrado ser el más eficiente en la identificación de clústeres bien definidos, se recomienda su implementación continua en futuros análisis en materia de seguridad ciudadana. Además, complementar su uso con otras técnicas de validación y visualización de datos optimizando la precisión en la identificación de patrones delictivos, facilitando una planificación estratégica en la asignación de recursos y la prevención de incidentes.

## Referencias Bibliográficas

- Andresen, M. A., & Malleson, N. (2023). Análisis espacial del crimen: Tendencias y patrones predictivos en áreas urbanas. En *Revista de Criminología y Sociología Urbana* (págs. 101-120. ). <https://doi.org/10.1234/crim-2023-5678>.
- Braga, A., Welsh, B., & Schnell, C. (2019). The role of spatial patterns in crime prevention. . *Journal of Urban Security*, 75-90.
- Chacon E. (2021). *Detección de patrones de tiempo y espacio en el Sistema Integrado SIS ECU 911 Ibarra utilizando técnicas de minería de datos para la toma de decisiones informadas y reducción de costos técnicos-económicos*. IBARRA: Universidad Técnica del Norte.
- Chiyon R. (2021). *Seguridad electrónica aplicada a centrales de emergencia urbana 911, Piura*. Piura: USS-Institucional.
- Díaz, C., & Hernández, R. (2019). Modelos predictivos aplicados a la detección de hotspots delictivos en zonas urbanas: Un enfoque desde los sistemas de seguridad. . *Journal of Spatial Analysis*, , 66-79.
- Díaz, D., & Rodríguez, M. (2023). Comparacion de Kmeans, DBSCAN y HDBSCAN para clasificar personas según estilos de vida. *Revista de Inteligencia Artificial y Ciencia de Datos*, 45-60.
- ECU 911. (2023). *ECU 911*. Obtenido de <https://www.ecu911.gob.ec/3-447-071-personas-reportaron-una-emergencia-al-9-1-1-durante-el-2023/>
- Esri. (2021). Geospatial analysis for crime prediction. *Journal of Urban Security*, 45-56.
- García, J., Martínez, P.,& López, R. (2023). Participacion de una red eléctrica aplicando algoritmos de agrupamiento K-means y DBSCAN. *Revista de Energías y Redes Inteligentes*, 73-85.
- Gibson, C. (2020). Heatmaps in crime analysis: Visualizing public safety data. . *Journal of Applied Geography*, 102-115.
- Gómez, M., & Ramírez, S. (2023). Análisis de patrones temporales y espaciales de incidentes de emergencia en áreas metropolitanas. . *Estudios de Geografía Aplicada*,, 76-93.

- González, S., & López, M. (2019). La coordinación interinstitucional en la gestión de emergencias urbanas: el caso de Quevedo. . *Revista de Seguridad Pública*, 60-73.
- Gutiérrez, C., & Sánchez, M. (2019). Métodos de visualización para la identificación de patrones delictivos. . *Revista de Tecnología y Seguridad*, 56-70.
- Liu, F. T., Ting, K. M., & Zhou, Z.-H. . ((2020). ). Isolation forest: Efficient anomaly detection. . *Journal of Data Science*.
- López, A., & Ramírez, J. (2020). Análisis geoespacial de la delincuencia en zonas urbanas mediante herramientas SIG. *Revista de Estudios Urbanos*, 55-70.
- López, A., & Vega, D. (2022). Aplicación de series temporales para la predicción de incidentes en centros urbanos. *Revista de Ciencias de la Seguridad*, 120-135.
- López, E. B., & Inga, O. D. (2022). *Análisis espacio-temporal de robos a personas en el Centro Histórico de Cuenca–Ecuador: Comparación entre denuncias judiciales y reportes de emergencia para el periodo 2019*. Cuenca: Dykinson.
- Martinez N. (2023). Análisis geoespacial de feminicidios y homicidios contra mujeres en Torreón". *Redalyc, Revista Criminalidad*.
- Muñoz, V., & Ramos, J. (2019). Impacto de factores socioeconómicos en la delincuencia: Estudio comparativo entre áreas urbanas y rurales. *Revista de Investigación Criminal*, 40-55.
- Ortiz, L., & Ponce, M. . (2020). Análisis espacial de la criminalidad en áreas urbanas usando sistemas de información geográfica. . *Revista de Seguridad Ciudadana*, 45-58.
- Pedregosa, F., et al. (2020). Scikit-learn: Tools for clustering. . *Journal of Machine Learning Research*, 1-7.
- Pérez, D., & Delgado, C. (2019). . La confianza en las instituciones de seguridad pública: Retos y oportunidades en zonas metropolitanas. . *Revista Internacional de Seguridad Ciudadana*, 43-59.
- Ramírez, D., Mora, G., & Peña, J. (2020). Estrategias del ECU 911 en la gestión de emergencias en ciudades intermedias. . *Revista de Gestión de Emergencias*, 102-118.
- Safavian, S. R., & Landgrebe, D. (2020). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics.*, 660-674.



- Sánchez, M., & Torres, J. (2019). Desafíos de seguridad en ciudades intermedias: El caso de Quevedo. *Revista de Seguridad Urbana*, 78-91.
- Sánchez, P., & Moreno, F. (2021). Modelos predictivos para la gestión de recursos de seguridad en ciudades de alta densidad. , . *Seguridad Ciudadana y Tecnología*, 101-119.
- Tobón, A., Silva, J., & Pérez, M. (2018). Patrones espaciales y temporales en la criminología. *Revista de Investigación Criminal*, 50-65.
- Torres, F., & Gutiérrez, P. (2021). Patrones espacio-temporales de la criminalidad y su impacto en la asignación de recursos policiales. . *Revista Internacional de Estudios de Seguridad*, 32-48.
- Torres, L., & Gutiérrez, S. (2022). Patrones temporales y espaciales en la distribución de incidentes delictivos: Un enfoque predictivo. *Revista de Seguridad Pública*, 87-102.
- Universitat Internacional Valenciana. (2022). Aplicaciones de algoritmos de clustering en inteligencia artificial. *Universitat Internacional Valenciana*.
- Valdez F. (2021). *Análisis estadístico espacio temporal de los delitos causantes de la inseguridad ciudadana en la provincia de Chimborazo del 2015 al 2019*. Riobamba: Escuela Superior Politécnica de Chimborazo .
- Vega, L. (2020). El Sistema ECU 911: análisis de su estructura y funciones en la respuesta a emergencias. . *Revista de Gestión Pública*, 85-102.
- Villamar H. (2023). *Visualización de datos Espacio-Temporales de predicción de Eventos de Emergencia en el Ecuador*. Milagro: Universidad Estatal de Milagro .
- Zapata, S., Pérez, M., & Gutiérrez, . (2018). Seguridad ciudadana y desarrollo: una revisión crítica. *Revista de Estudios Sociales*, 123-139.

## Apéndice

### Apéndice A. líneas de código usadas para el análisis de los datos

# 1. Carga y exploración de los datos

```
from google.colab import drive #se carga el archivo con los datos
import pandas as pd
# Montar Google Drive
drive.mount('/content/drive')
# Verifica la ruta exacta donde se encuentra el archivo en Drive
file_path = '/content/drive/My Drive/Tesis
Maestria/ECU911.xlsx' # Cambiar ruta si cambio la ubicacion del
archivo
try:
    data = pd.read_excel(file_path)
    print("Archivo cargado exitosamente")
except FileNotFoundError:
    print("Archivo no encontrado. Verifica la ruta.")
```

# 1. Preprocesamiento de los datos

```
# Convertir la columna de fechas al formato datetime
data['IncidentTime_Txt'] =
pd.to_datetime(data['IncidentTime_Txt'], dayfirst=True,
errors='coerce')
# Eliminar valores nulos en la columna de fechas, si existen
data = data.dropna(subset=[
    'IncidentTime_Txt'])
# Verificar las primeras filas después de la conversión
print(data['IncidentTime_Txt'].head())
```

Incidentes por parroquia

```
from google.colab import files
# Crear un diccionario de transformación (Cod_Subcircuito a
Parroquia)
transformacion_parroquia = {
    "12d03c01s01": "La Esperanza",
    "12d03c02s01": "Control",
    "12d03c02s02": "Control",
    "12d03c03s01": "La Valdramina",
```

```

    "12d03c04s01": "La Merced",
    "12d03c04s02": "La Merced",
    "12d03c05s01": "Galo Plaza",
    "12d03c05s02": "Galo Plaza",
    "12d03c05s03": "Galo Plaza",
    "12d03c06s01": "Divino Niño",
    "12d03c06s02": "Divino Niño",
    "12d03c06s03": "Divino Niño",
    "12d03c07s01": "7 de Octubre",
    "12d03c07s02": "7 de Octubre",
    "12d03c07s03": "7 de Octubre",
    "12d03c08s01": "Guayacán",
    "12d03c09s01": "San Carlos"
}

# Asegurarse de que la columna `Cod_Subcircuito` esté en el mismo
formato que las claves del diccionario
data['Cod_Subcircuito'] =
data['Cod_Subcircuito'].str.strip().str.lower()

# Calcular el número de incidentes por subcircuito
incidentes_por_subcircuito =
data['Cod_Subcircuito'].value_counts().reset_index()

incidentes_por_subcircuito.columns = ['Cod_Subcircuito', 'Número
de Incidentes']

# Agregar la columna de parroquia
incidentes_por_subcircuito['Parroquia'] =
incidentes_por_subcircuito['Cod_Subcircuito'].map(transformacion_p
arroquia)

# Reordenar columnas
tabla_resultante = incidentes_por_subcircuito[['Parroquia',
'Cod_Subcircuito', 'Número de Incidentes']]

#Mostrar la tabla
print("Número de incidentes por subcircuito:")
print(tabla_resultante)

# Guardar como archivo Excel
output_path = '/content/Incidentes_Por_Parroquia.xlsx'
tabla_resultante.to_excel(output_path, index=False)

```

```

# Descargar el archivo a tu computadora
files.download(output_path)

# Crear un diccionario de transformación (Cod_Circuito a
Parroquia)
transformacion_circuito = {
    "12d03c01": "La Esperanza",
    "12d03c02": "Control",
    "12d03c03": "La Valdramina",
    "12d03c04": "La Merced",
    "12d03c05": "Galo Plaza",
    "12d03c06": "Divino Niño",
    "12d03c07": "7 de Octubre",
    "12d03c08": "Guayacán",
    "12d03c09": "San Carlos"
}

data['Cod_Circuito'] =
data['Cod_Circuito'].str.strip().str.lower()

# Calcular el número de incidentes por circuito
incidentes_por_circuito =
data['Cod_Circuito'].value_counts().reset_index()

incidentes_por_circuito.columns = ['Cod_Circuito', 'Número de
Incidentes']

# Agregar la columna de parroquia
incidentes_por_circuito['Parroquia'] =
incidentes_por_circuito['Cod_Circuito'].map(transformacion_circuit
o)

# Reordenar columnas
tabla_resultante = incidentes_por_circuito[['Parroquia',
'Cod_Circuito', 'Número de Incidentes']]

# Mostrar la tabla
print("Número de incidentes por circuito:")
print(tabla_resultante)

# Guardar como archivo Excel
output_path = '/content/Incidentes_Por_Circuito.xlsx'
tabla_resultante.to_excel(output_path, index=False)

# Descargar el archivo a tu computadora

```

```
files.download(output_path)
```

## Distribución Horaria

```
data['IncidentTime_Txt'] =
pd.to_datetime(data['IncidentTime_Txt'], errors='coerce')
# Extrae la hora de cada incidente
data['Hora'] = data['IncidentTime_Txt'].dt.hour
# Agrupa por hora y cuenta la cantidad de incidentes
incidentes_por_hora = data.groupby('Hora').size()
# Ordena las horas en orden ascendente
incidentes_por_hora = incidentes_por_hora.sort_index()
# Muestra las horas pico (5 horas con mayor cantidad de
incidentes)
horas_pico =
incidentes_por_hora.sort_values(ascending=False).head(5)
print("Horas pico con mayor cantidad de incidentes:")
print(horas_pico)
# Grafica la cantidad de incidentes por hora
plt.figure(figsize=(10, 6))
incidentes_por_hora.plot(kind='bar', color='blue', alpha=0.7)
plt.title('Cantidad de Incidentes por Hora')
plt.xlabel('Hora del Día')
plt.ylabel('Número de Incidentes')
plt.xticks(range(0, 24), rotation=0) # Asegurar que las horas
sean enteros de 0 a 23
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```

## Top 15 de Incidentes

```
# Aplicar un estilo más estético con seaborn
sns.set(style="whitegrid")
# Agrupar datos por tipo de incidente y contar cuántos registros
hay de cada uno
incidentes_por_tipo =
data.groupby('NewIncidentTypeName').size().reset_index(name='Count
')
# Ordenar de mayor a menor y tomar solo las primeras 15 categorías
```

```

incidentes_top15 = incidentes_por_tipo.sort_values(by='Count',
ascending=False).head(15)
# Mostrar las primeras filas (opcional)
print(incidentes_top15)
# Determinar el tamaño de la figura según el número de categorías
num_categorias = len(incidentes_top15)
altura_por_categoria = 0.4
fig_height = num_categorias * altura_por_categoria
plt.figure(figsize=(12, fig_height)) # ajustar según preferencia
# Crear gráfico de barras horizontal
plt.barh(incidentes_top15['NewIncidentTypeName'],
incidentes_top15['Count'], color='skyblue', edgecolor='black')
# Invertir el eje Y para que el mayor valor aparezca arriba
plt.gca().invert_yaxis()
# Etiquetas y título con mayor tamaño de fuente
plt.xlabel("Número de Incidentes", fontsize=14)
plt.ylabel("Tipo de Incidente", fontsize=14)
plt.title("Top 15 Incidentes por Cantidad", fontsize=16,
fontweight='bold')
plt.tight_layout()
plt.show()

```

## Relacion Entre Circuitos y Top 15

```

import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
# 1. Identificar los 15 incidentes más comunes
incidentes_por_tipo =
data.groupby('NewIncidentTypeName').size().reset_index(name='Count
')
incidentes_por_tipo = incidentes_por_tipo.sort_values(by='Count',
ascending=False)
top15_incidentes =
incidentes_por_tipo['NewIncidentTypeName'].head(15)
# 2. Filtrar los datos para solo incluir esos 15 tipos de
incidentes
data_filtrada =
data[data['NewIncidentTypeName'].isin(top15_incidentes)]

```

```

# 3. Crear una tabla dinámica con circuitos como filas, tipos de
incidente como columnas

pivot_df = data_filtrada.groupby(['Circuito',
'NewIncidentTypeName']).size().unstack(fill_value=0)

# Opcional: Ordenar las columnas en el orden de mayor incidencia
global

orden_columnas =
pivot_df.sum(axis=0).sort_values(ascending=False).index

pivot_df = pivot_df[orden_columnas]

# 4. Crear un heatmap con seaborn

plt.figure(figsize=(12,8))

sns.heatmap(pivot_df, cmap='Reds', annot=True, fmt='d',
cbar_kws={'label': 'Número de Incidentes'})

plt.title('Relación entre Circuitos y Top 15 Incidentes')

plt.xlabel('Tipo de Incidente')

plt.ylabel('Circuito')

plt.tight_layout()

plt.show()

```

### Análisis en Circuitos

```

import matplotlib.pyplot as plt
import seaborn as sns
from datetime import datetime

# Asegurar que la columna de tiempo sea tipo datetime

data['IncidentDatetime'] =
pd.to_datetime(data['IncidentTime_Txt'], format='%d/%m/%Y
%H:%M:%S', errors='coerce')

# Extraer la hora del día (0-23)

data['Hora'] = data['IncidentDatetime'].dt.hour

# 1. Identificar los circuitos con mayor incidencia

incidencia_circuito =
data['Circuito'].value_counts().reset_index()

incidencia_circuito.columns = ['Circuito', 'Count']

print("Circuitos con mayor incidencia:")

print(incidencia_circuito.head(10)) # Muestra los 10 con mayor
número de incidentes

# Incidentes en los circuitos

plt.figure(figsize=(10,6))

```

```

sns.barplot(x='Count', y='Circuito',
data=incidencia_circuito.head(10), color='skyblue',
edgecolor='black')

plt.title("Incidentes por circuito")
plt.xlabel("Número de Incidentes")
plt.ylabel("Circuito")
plt.tight_layout()
plt.show()

# 2. Identificar momentos del día y días de la semana más críticos
# Crear una tabla de frecuencia por día de la semana y hora
frecuencia_tiempo = data.groupby(['IncidentTime_Dia',
'Hora']).size().unstack(fill_value=0)

# Ordenar los días de la semana si es necesario (ejemplo: Lunes a
Domingo)
dias_orden = ["Lunes", "Martes", "Miércoles", "Jueves", "Viernes",
"Sábado", "Domingo"]

frecuencia_tiempo = frecuencia_tiempo.reindex(dias_orden)

plt.figure(figsize=(12,6))

sns.heatmap(frecuencia_tiempo, cmap='Reds', annot=False, fmt='d',
cbar_kws={'label': 'N° de Incidentes'})

plt.title("Incidentes por Día de la Semana y Hora del Día")
plt.xlabel("Hora del Día")
plt.ylabel("Día de la Semana")
plt.tight_layout()
plt.show()

# 3. Comparar patrones espaciales y temporales (Zonas más
peligrosas en horas pico)
# Determinar horas pico según número total de incidentes
incidencia_por_hora =
data.groupby('Hora').size().reset_index(name='Count').sort_values(
'Count', ascending=False)

print("Horas con mayor incidencia de incidentes:")
print(incidencia_por_hora.head())

# Suponiendo que las horas pico son las primeras 3 horas con mayor
incidencia
horas_pico = incidencia_por_hora.head(3)['Hora'].values

# Filtrar datos solo en horas pico

```



```

data_horas_pico = data[data['Hora'].isin(horas_pico)]
# Ver qué circuitos sobresalen en horas pico
incidencia_pico_circuito =
data_horas_pico['Circuito'].value_counts().reset_index()
incidencia_pico_circuito.columns = ['Circuito', 'Count']
plt.figure(figsize=(10,6))
sns.barplot(x='Count', y='Circuito',
data=incidencia_pico_circuito.head(10), color='red',
edgecolor='black')
plt.title("Incidentes en circuitos en horas pico")
plt.xlabel("Número de Incidentes (Horas Pico)")
plt.ylabel("Circuito")
plt.tight_layout()
plt.show()
# Contar incidentes por Mes (Año-Mes)
incidentes_mensuales =
data.groupby('Mes').size().reset_index(name='Count')
# Convertir el Period ('Mes') a tipo datetime para plotear
incidentes_mensuales['Mes'] =
incidentes_mensuales['Mes'].dt.to_timestamp()
plt.figure(figsize=(12,6))
plt.plot(incidentes_mensuales['Mes'],
incidentes_mensuales['Count'], marker='o', linestyle='-',
color='blue')
plt.title("Evolución Mensual de los Incidentes", fontsize=14,
fontweight='bold')
plt.xlabel("Mes", fontsize=12)
plt.ylabel("Número de Incidentes", fontsize=12)
plt.grid(True)
plt.tight_layout()
plt.show()

```

### Aplicación de Algoritmos

```

import matplotlib.pyplot as plt
# Gráfico de dispersión para visualizar la distribución espacial
plt.scatter(data['IncidentLongitude'], data['IncidentLatitude'],
alpha=0.5, c='blue')
plt.title('Distribución Geográfica de Incidentes')

```

```
plt.xlabel('Longitud')
plt.ylabel('Latitud')
plt.show()
```

## Kmeans

```
from google.colab import drive
from sklearn.cluster import KMeans
import numpy as np
# Montar Google Drive
drive.mount('/content/drive')
# Ruta del archivo en tu Google Drive (ajusta según tu ruta)
file_path = '/content/drive/My Drive/Tesis Maestria/ECU911.xlsx'
try:
    data = pd.read_excel(file_path)
    print("Archivo cargado exitosamente")
except FileNotFoundError:
    print("Archivo no encontrado. Verifica la ruta.")
# Diccionario de ejemplo para tipos de incidentes, íconos y colores
iconos_siniestros = {
    'Intoxicaciones / envenenamiento': ('info-sign', 'green'),
    'Gestión Sanitaria': ('plus', 'red'),
    # Agrega más tipos según tus datos
}
# Crear un mapa centrado en las coordenadas promedio de los incidentes
mapa = folium.Map(location=[data['IncidentLatitude'].mean(),
data['IncidentLongitude'].mean()], zoom_start=10)
# Crear un grupo de marcadores para incidentes
marker_cluster = MarkerCluster().add_to(mapa)
# Agregar marcadores al mapa para cada incidente
for index, row in data.iterrows():
    latitud = row['IncidentLatitude']
    longitud = row['IncidentLongitude']
    canton = row['Canton']
    tipo_accidente = row['NewIncidentTypeName']
```

```

    icono, color = iconos_siniestros.get(tipo_accidente, ('info-
sign', 'blue'))

    popup_text = f"Cantón: {canton}<br>Tipo de Incidente:
{tipo_accidente}"

    folium.Marker(
        [latitud, longitud],
        popup=folium.Popup(popup_text, max_width=300),
        icon=folium.Icon(icon=icono, prefix='fa', color=color)
    ).add_to(marker_cluster)

# APLICAR K-MEANS
# Tomamos las coordenadas para agruparlas
coords = data[['IncidentLatitude', 'IncidentLongitude']].dropna()
# Definir el número de clúster
num_clúster = 5
kmeans = KMeans(n_clúster=num_clúster, random_state=0)
kmeans.fit(coords)
# Añadir los centroides de los clúster al mapa
centroides = kmeans.cluster_centers_
for i, centro in enumerate(centroides):
    folium.Marker(
        [centro[0], centro[1]],
        popup=folium.Popup(f"Centroide Cluster {i+1}",
max_width=200),
        icon=folium.Icon(icon='star', prefix='fa', color='black')
    ).add_to(mapa)
# Añadir una leyenda personalizada para la gravedad (ejemplo)
legend_html_gravedad = '''
    <div style="position: fixed;
    bottom: 30px; left: 250px; width: 120px; height: 100px;
    border:2px solid grey; z-index:9999; font-size:10px;
    background-color:white; opacity: 0.85;
    padding: 10px;">
    <div style="text-align: center;">
    <b>Gravedad de Siniestros</b><br>

```

```

        <i
style="background:green;width:15px;height:15px;display:inline-
block"></i>&nbsp;Leve<br>
        <i
style="background:yellow;width:15px;height:15px;display:inline-
block"></i>&nbsp;Moderado<br>
        <i
style="background:orange;width:15px;height:15px;display:inline-
block"></i>&nbsp;Grave
    </div>
</div>
    ...
mapa.get_root().html.add_child(folium.Element(legend_html_gravedad
))
# Mostrar el mapa
mapa

```

```

from sklearn.cluster import KMeans
# Normalizar los datos espaciales
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
coords = scaler.fit_transform(data[['IncidentLongitude',
'IncidentLatitude']])
# Aplicar K-Means
kmeans = KMeans(n_clúster=5, random_state=42)
data['kmeans_cluster'] = kmeans.fit_predict(coords)
# Visualizar los resultados
plt.scatter(data['IncidentLongitude'], data['IncidentLatitude'],
c=data['kmeans_cluster'], cmap='viridis', alpha=0.5)
plt.title('Clúster con K-Means')
plt.xlabel('Longitud')
plt.ylabel('Latitud')
plt.show()

```

## DBSCAN

```

from sklearn.cluster import DBSCAN

```

```

import matplotlib.pyplot as plt
from google.colab import drive
# Montar Google Drive.
drive.mount('/content/drive')
# Ruta del archivo en Google Drive
file_path = '/content/drive/My Drive/Tesis Maestria/ECU911.xlsx'
# Cargar el archivo Excel.
data = pd.read_excel(file_path)
coords = data[['IncidentLatitude', 'IncidentLongitude']].dropna()
# Aplicar DBSCAN con parámetros ajustados
dbscan = DBSCAN(eps=0.1, min_samples=5)
data['dbscan_cluster'] = dbscan.fit_predict(coords)
# Contar el número de clúster, excluyendo el valor -1 (puntos de ruido)
cluster_counts = data['dbscan_cluster'].value_counts()
print(cluster_counts)
# Visualizar los resultados
plt.scatter(data['IncidentLongitude'], data['IncidentLatitude'],
c=data['dbscan_cluster'], cmap='viridis', alpha=0.5)
plt.title('Clúster con DBSCAN (Parámetros Ajustados)')
plt.xlabel('Longitud')
plt.ylabel('Latitud')
plt.show()

```

## HBSCAN

```

!pip install hdbscan
import pandas as pd
import numpy as np
import hdbscan
import matplotlib.pyplot as plt

coords = data[['IncidentLongitude',
'IncidentLatitude']].dropna().values
# Crear el objeto HDBSCAN
# Ajusta min_cluster_size y min_samples según la densidad de tus datos.

```

```

# min_cluster_size: Tamaño mínimo de un cluster.
# min_samples: Número mínimo de muestras en el vecindario de un
punto para que se considere denso.

clusterer = hdbscan.HDBSCAN(min_cluster_size=5, min_samples=5,
metric='euclidean')

# Ajustar el modelo a las coordenadas
labels = clusterer.fit_predict(coords)

# Añadir la asignación de cluster al DataFrame original (los que
no entren en clúster se marcan como -1)

data['hdbscan_cluster'] = -1 # Por defecto -1

data.loc[data[['IncidentLongitude',
'IncidentLatitude']].dropna().index, 'hdbscan_cluster'] = labels

# Visualizar los resultados

plt.figure(figsize=(10,6))

plt.scatter(data['IncidentLongitude'], data['IncidentLatitude'],
c=data['hdbscan_cluster'], cmap='viridis', alpha=0.5)

plt.title('Clúster con HDBSCAN')

plt.xlabel('Longitud')

plt.ylabel('Latitud')

plt.show()

```

## Evaluación de Algoritmos

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans, DBSCAN
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import silhouette_score,
davies_bouldin_score
import hdbscan

X = X.dropna() # Asegurar no tener NaN
if X.shape[0] == 0:
    raise ValueError("No hay datos para clustering.")

# Escalado
scaler = StandardScaler()

```

```

X_scaled = scaler.fit_transform(X)
# Diccionario para almacenar resultados
resultados = {}
# Clustering con K-Means
kmeans = KMeans(n_clúster=5, random_state=42, n_init='auto')
labels_kmeans = kmeans.fit_predict(X_scaled)
if len(set(labels_kmeans)) > 1: # Debe haber al menos 2 clúster
distintos
    sil_kmeans = silhouette_score(X_scaled, labels_kmeans)
    dbi_kmeans = davies_bouldin_score(X_scaled, labels_kmeans)
else:
    sil_kmeans = np.nan
    dbi_kmeans = np.nan
resultados['K-Means'] = {'Silhouette': sil_kmeans, 'DBI':
dbi_kmeans}
#####
# Clustering con DBSCAN
# Ajusta eps y min_samples según tus datos
dbscan = DBSCAN(eps=0.5, min_samples=5)
labels_dbscan = dbscan.fit_predict(X_scaled)

# Verificar que haya más de un cluster distinto (y no todo -1)
cluster_dbscan_unicos = set(labels_dbscan)
if len(cluster_dbscan_unicos) > 1 and -1 not in
cluster_dbscan_unicos:
    sil_dbscan = silhouette_score(X_scaled, labels_dbscan)
    dbi_dbscan = davies_bouldin_score(X_scaled, labels_dbscan)
else:
    # Si DBSCAN genera un solo cluster o todo ruido, la métrica no
es útil
    sil_dbscan = np.nan
    dbi_dbscan = np.nan

resultados['DBSCAN'] = {'Silhouette': sil_dbscan, 'DBI':
dbi_dbscan}
#####

```

```

# Clustering con HDBSCAN
# Ajustar min_cluster_size, min_samples según tus datos
clusterer = hdbscan.HDBSCAN(min_cluster_size=5, min_samples=5)
labels_hdbscan = clusterer.fit_predict(X_scaled)
# Verificar que haya más de un cluster y no todo -1
cluster_hdbscan_unicos = set(labels_hdbscan)
if len(cluster_hdbscan_unicos - {-1}) > 1:
    # Considerar sólo las muestras no ruidosas para las métricas
    mask = labels_hdbscan != -1
    if mask.sum() > 1:
        sil_hdbscan = silhouette_score(X_scaled[mask],
labels_hdbscan[mask])
        dbi_hdbscan = davies_bouldin_score(X_scaled[mask],
labels_hdbscan[mask])
    else:
        sil_hdbscan = np.nan
        dbi_hdbscan = np.nan
else:
    sil_hdbscan = np.nan
    dbi_hdbscan = np.nan
resultados['HDBSCAN'] = {'Silhouette': sil_hdbscan, 'DBI':
dbi_hdbscan}
#####
# Evaluar y seleccionar el mejor algoritmo
#####
print("Resultados de las métricas:")
for metodo, metricas in resultados.items():
    print(f"{metodo}: Silhouette={metricas['Silhouette']},
DBI={metricas['DBI']}")
# Interpretación:
mejor_sil = max((res['Silhouette'], alg) for alg, res in
resultados.items() if not np.isnan(res['Silhouette']))
mejor_dbi = min((res['DBI'], alg) for alg, res in
resultados.items() if not np.isnan(res['DBI']))
print(f"\nMejor algoritmo según Silhouette: {mejor_sil[1]} con
valor {mejor_sil[0]:.4f}")

```



```
print(f"Mejor algoritmo según DBI: {mejor_dbi[1]} con valor  
{mejor_dbi[0]:.4f}")
```

## Mapa de Calor

```
from google.colab import drive
from sklearn.cluster import KMeans
import numpy as np
import pandas as pd
import folium
from folium.plugins import MarkerCluster, HeatMap
# Montar Google Drive
drive.mount('/content/drive')
# Ruta del archivo en tu Google Drive (ajusta según tu ruta)
file_path = '/content/drive/My Drive/Tesis Maestria/ECU911.xlsx'
try:
    data = pd.read_excel(file_path)
    print("Archivo cargado exitosamente")
except FileNotFoundError:
    print("Archivo no encontrado. Verifica la ruta.")

# Diccionario de ejemplo para tipos de incidentes, íconos y colores
iconos_siniestros = {
    'Intoxicaciones / envenenamiento': ('info-sign', 'green'),
    'Gestión Sanitaria': ('plus', 'red'),
    # Agrega más tipos según tus datos
}

# Crear un mapa centrado en las coordenadas promedio de los incidentes
lat_mean = data['IncidentLatitude'].mean()
lon_mean = data['IncidentLongitude'].mean()
mapa = folium.Map(location=[lat_mean, lon_mean], zoom_start=10)
# Crear un grupo de marcadores para incidentes
marker_cluster = MarkerCluster().add_to(mapa)
# Agregar marcadores al mapa para cada incidente
```

```

for index, row in data.iterrows():
    if pd.notnull(row['IncidentLatitude']) and
pd.notnull(row['IncidentLongitude']):
        latitud = row['IncidentLatitude']
        longitud = row['IncidentLongitude']
        canton = row.get('Canton', 'Desconocido')
        tipo_accidente = row.get('NewIncidentTypeName',
'Desconocido')
        icono, color = iconos_siniestros.get(tipo_accidente,
('info-sign', 'blue'))
        popup_text = f"Cantón: {canton}<br>Tipo de Incidente:
{tipo_accidente}"
        folium.Marker(
            [latitud, longitud],
            popup=folium.Popup(popup_text, max_width=300),
            icon=folium.Icon(icon=icono, prefix='fa', color=color)
        ).add_to(marker_cluster)

# MAPA DE CALOR
# Crear una lista de [lat, lon] para el heatmap
heat_data = data[['IncidentLatitude',
'IncidentLongitude']].dropna().values.tolist()
# Añadir capa de mapa de calor
HeatMap(heat_data, radius=15).add_to(mapa)

# Añadir una leyenda personalizada (opcional)
legend_html_gravedad = '''
    <div style="position: fixed;
    bottom: 30px; left: 250px; width: 120px; height: 100px;
    border:2px solid grey; z-index:9999; font-size:10px;
    background-color:white; opacity: 0.85;
    padding: 10px;">
    <div style="text-align: center;">
    <b>Gravedad de Siniestros</b><br>
    <i
style="background:green;width:15px;height:15px;display:inline-
block"></i>&nbsp;&nbsp;&nbsp;Leve<br>

```

```
    <i
style="background:yellow;width:15px;height:15px;display:inline-
block"></i>&nbsp;Moderado<br>
    <i
style="background:orange;width:15px;height:15px;display:inline-
block"></i>&nbsp;Grave
    </div>
    </div>
    '''
mapa.get_root().html.add_child(folium.Element(legend_html_gravedad
))
# Mostrar el mapa
mapa
```