

**Análisis y visualización de desastres naturales en Colombia: identificación de patrones  
para la mejora de sistemas de alerta y planificación de contingencias**

Claudia Ximena Saza Baraceta

Jefferson Daniel Munevar Herrera

Asesor

Rafael Gaitán Ospina

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Especialización en Ciencia de Datos y Analítica

2025

## Resumen

Este proyecto de grado se centra en el desarrollo de una plataforma interactiva de visualización de datos utilizando Power BI para analizar y comprender patrones históricos de desastres naturales en Colombia. Mediante el análisis de eventos como terremotos, incendios e inundaciones, se busca identificar tendencias recurrentes que permitan mejorar los sistemas de alerta y planificación de contingencias. Se emplearon técnicas avanzadas de análisis de datos, como la segmentación mediante clústeres, para categorizar zonas de riesgo y optimizar estrategias de mitigación.

Entre los hallazgos más relevantes se destaca que las inundaciones son el evento con mayor impacto en infraestructura, mientras que los incendios forestales son los más frecuentes. Asimismo, los departamentos de Cundinamarca y Antioquia concentran el mayor número de eventos. El análisis trimestral y por clúster permitió identificar patrones estacionales y zonas con vulnerabilidades específicas, lo cual aporta insumos valiosos para la gestión del riesgo a nivel territorial. La plataforma desarrollada proporciona herramientas de visualización dinámica que apoyan la toma de decisiones de entidades gubernamentales y organismos de atención de emergencias.

***Palabras clave:*** Visualización, datos, desastres naturales, análisis de patrones, planificación de contingencias, clústeres.

## Abstract

This thesis focuses on the development of an interactive data visualization platform using Power BI to analyze and understand historical patterns of natural disasters in Colombia. By analyzing events such as earthquakes, fires, and floods, the project seeks to identify recurring trends that will improve warning systems and contingency planning. Advanced data analysis techniques, such as cluster segmentation, were used to categorize risk zones and optimize mitigation strategies.

Among the most relevant findings, it is worth noting that floods are the event with the greatest impact on infrastructure, while forest fires are the most frequent. Furthermore, the departments of Cundinamarca and Antioquia concentrate the highest number of events. Quarterly and cluster analysis allowed for the identification of seasonal patterns and areas with specific vulnerabilities, which provides valuable input for risk management at the territorial level. The developed platform provides dynamic visualization tools that support decision-making by government entities and emergency response agencies.

**Keywords:** Visualization, data, natural disasters, pattern analysis, contingency planning, clusters.

## Tabla de Contenido

Introducción .....	11
Planteamiento del Problema .....	13
Justificación .....	14
Objetivos .....	15
Objetivo General .....	15
Objetivos Específicos.....	15
Marco Teórico.....	16
Análisis de Datos en Power BI .....	16
Ciencias del Análisis de Datos y Minería de Datos .....	16
Segmentación de Datos y Análisis de Clústeres .....	17
Visualización de Datos e Inteligencia de Negocios .....	17
Zonas de Amenaza y Riesgo.....	18
Metodología .....	20
Recolección y Preparación de Datos.....	20
Fuentes de Datos .....	20
Variables para Analizar .....	20
Procesamiento Inicial .....	20
Consolidación de la Información y Segmentación de Datos .....	20
Análisis Exploratorio.....	20
Técnicas Analíticas.....	20
Interpretación de Resultados .....	20
Diseño y Visualización en Power BI .....	21

Herramienta Principal.....	21
Funcionalidades Clave .....	21
Pruebas de Usabilidad .....	21
Evaluación y Conclusiones .....	21
Comparación de Enfoques.....	21
Retroalimentación .....	21
Documentación.....	21
Limpieza de Bases de Datos .....	22
Limpieza de las Bases de Datos Previo a la Consolidación.....	22
Visualización en Power BI.....	25
Análisis de Datos y Graficas.....	28
Análisis Descriptivo .....	28
Impactos y Afectaciones .....	39
Heridos, Fallecidos, Desaparecidos.....	39
Daños a Infraestructura (Puentes, Vías, Viviendas).....	41
Análisis de Clúster .....	45
Características de Cada Grupo .....	46
Clúster 0 – Incendios Forestales como Evento Dominante.....	46
Clúster 1–Inundaciones y Movimientos en Masa en Antioquia.....	48
Clúster 2 – Movimiento en Masa en el Suroccidente y Oriente.....	49
Clúster 3 – Vendavales y su Concentración en la Región Caribe y Eje Cafetero .....	51
Clúster 4 – Deslizamientos y Creciente Súbita en Zonas con Relieve.....	53
Comparaciones entre Clústeres .....	55

Tipo de Desastre Dominante .....	56
Ubicación Geográfica Predominante.....	57
Potencial de Impacto y Peligrosidad .....	58
Complejidad para la Gestión del Riesgo .....	59
Análisis Comparativo con la NSR-10 .....	60
Plan para la Optimización de Medidas de Mitigación y Respuesta ante Desastres Naturales .....	63
Fortalecimiento de Sistemas de Monitoreo y Alerta Temprana.....	63
Planeación Territorial Basada en Datos .....	63
Educación Comunitaria y Gestión Local del Riesgo .....	63
Optimización de Recursos para Respuesta Rápida .....	64
Evaluación y Retroalimentación Continúa Basada en Visualización .....	64
Conclusiones .....	65
Referencias.....	67
Apéndices.....	69

## Lista de Tablas

<b>Tabla 1</b> <i>Consolidado Comparativo de los Indicadores Anuales</i> .....	29
<b>Tabla 2</b> <i>Consolidado por Año el Número Total de Heridos, Fallecidos y Desaparecidos por Tipo de Evento</i> .....	40
<b>Tabla 3</b> <i>Consolidado Años por Evento con Mayo Total de Daños en Infraestructura</i> .....	43
<b>Tabla 4</b> <i>Tipo de Desastre Predominante por cada Clúster</i> .....	56
<b>Tabla 5</b> <i>Ubicación Geográfica Predominante por Clúster</i> .....	58
<b>Tabla 6</b> <i>Potencial de Impacto y Peligrosidad por Clúster</i> .....	59
<b>Tabla 7</b> <i>Complejidad para la Gestión del Riesgo por Clúster</i> .....	60

## Lista de Figuras

<b>Figura 1</b> <i>Mapa de Zonas de Amenaza Sísmica NSR-10</i> .....	19
<b>Figura 2</b> <i>Visualización Resumen Geográfico de Eventos</i> .....	26
<b>Figura 3</b> <i>Visualización Eventos por Trimestre: Distribución y Proporción Según Tipo de Desastre</i> .....	27
<b>Figura 4</b> <i>Análisis De Datos Predominantes En Base De Datos</i> .....	29
<b>Figura 5</b> <i>Total de Eventos por Año</i> .....	31
<b>Figura 6</b> <i>Eventos por Año y Mes para Cundinamarca vs Incendio Forestal</i> .....	32
<b>Figura 7</b> <i>Eventos por Año y Mes para Antioquia vs Movimiento en Masa</i> .....	34
<b>Figura 8</b> <i>Evento por Trimestre para Cundimarca</i> .....	35
<b>Figura 9</b> <i>Evento por Trimestre para Antioquia</i> .....	37
<b>Figura 10</b> <i>Resumen General del Número Total de Heridos, Fallecidos y Desaparecidos por Tipo de Evento</i> .....	39
<b>Figura 11</b> <i>Resumen Daños en Puentes, Vías y Viviendas</i> .....	42
<b>Figura 12</b> <i>Visualización de Clúster 0</i> .....	46
<b>Figura 13</b> <i>Detalle de Clúster 0</i> .....	47
<b>Figura 14</b> <i>Visualización de Clúster 1</i> .....	48
<b>Figura 15</b> <i>Detalle de Clúster 1</i> .....	49
<b>Figura 16</b> <i>Visualización de Clúster 2</i> .....	50
<b>Figura 17</b> <i>Detalle de Clúster 2</i> .....	51
<b>Figura 18</b> <i>Visualización de Clúster 3</i> .....	52
<b>Figura 19</b> <i>Detalle de Clúster 3</i> .....	53
<b>Figura 20</b> <i>Visualización de Clúster 4</i> .....	54

**Figura 21** *Detalle de Clúster 4*..... 55

**Lista de Apéndices**

<b>Apéndice A</b> <i>Enlace Plataforma Power BI</i> .....	69
---	----

## Introducción

En la actualidad, los desastres naturales representan una de las principales amenazas para la seguridad y el bienestar de las poblaciones alrededor del mundo. Fenómenos como terremotos, inundaciones, incendios forestales y tormentas han dejado a su paso devastación, pérdida de vidas y consecuencias económicas significativas. Ante esta realidad, la necesidad de desarrollar herramientas eficaces para la gestión del riesgo de desastres se vuelve imperativa. El presente proyecto de grado se propone abordar esta problemática mediante el desarrollo de una plataforma interactiva de visualización de datos que permita analizar y comprender patrones históricos de desastres naturales.

La plataforma utilizará Power BI como su herramienta principal, la cual ofrece capacidades avanzadas para la representación gráfica y la manipulación de datos. A través de la visualización dinámica de datos recopilados, se buscará facilitar el análisis de eventos históricos significativos. El enfoque estará en identificar tendencias y patrones recurrentes que no solo contribuyan a una mejor comprensión de los desastres pasados, sino que también sean útiles para la formulación de estrategias de prevención ante futuros eventos catastróficos.

Uno de los objetivos clave de este proyecto es servir como una herramienta de apoyo para planificadores de contingencia, autoridades gubernamentales y organismos de gestión del riesgo. Al proporcionar acceso a análisis en tiempo real, la plataforma permitirá a estos actores tomar decisiones fundamentadas en datos, lo que es crucial para la adaptación de infraestructuras y la planificación urbana resiliente. La integración de análisis, junto con la elaboración de modelos estadísticos, proporcionará una visión más clara de los riesgos potenciales y facilitará la implementación de estrategias de mitigación efectivas.

Desde una perspectiva técnica, el proyecto abarcará una serie de etapas esenciales que incluyen la recopilación, limpieza y procesamiento de datos. Se obtendrán datos de diversas fuentes, como registros históricos locales y organismos oficiales.

En última instancia, esta plataforma no solo se limitará a ser una herramienta de análisis visual; su desarrollo contribuirá significativamente al avance del conocimiento en el ámbito de la gestión del riesgo de desastres y el ordenamiento territorial. Al fortalecer las capacidades de respuesta ante eventos catastróficos, se espera que el proyecto ayude a reducir el impacto social y económico de estos fenómenos, promoviendo una cultura de prevención y preparación que sea esencial para la seguridad de las comunidades. Con este enfoque integral, se busca establecer una base sólida para el desarrollo de políticas públicas y estrategias que protejan a las poblaciones más vulnerables frente a la creciente amenaza de desastres naturales.

### **Planteamiento del Problema**

¿Cómo puede la visualización del histórico de desastres naturales, basada en el análisis de patrones de eventos pasados, contribuir a la mejora de los sistemas de alerta y planes de contingencia para futuros desastres?

### **Justificación**

Dada la creciente frecuencia e impacto de los desastres naturales, es crucial contar con herramientas que faciliten la interpretación de datos históricos para mejorar la predicción y prevención de estos eventos. Mediante Power BI, se podrán analizar patrones en los datos, fortaleciendo la toma de decisiones en la gestión del riesgo.

## **Objetivos**

### **Objetivo General**

Diseñar una plataforma de análisis y visualización de datos en Power BI para identificar patrones relevantes en el comportamiento y la frecuencia de desastres naturales históricos, contribuyendo a la optimización de medidas de mitigación y respuesta.

### **Objetivos Específicos**

Analizar la frecuencia y el comportamiento de diversos tipos de desastres naturales utilizando datos históricos.

Identificar patrones recurrentes mediante técnicas avanzadas de análisis de datos.

Aplicar técnicas de análisis de clústeres para segmentar las regiones o tipos de desastres.

Diseñar e implementar una plataforma visual interactiva en Power BI para facilitar la toma de decisiones.

## **Marco Teórico**

El marco teórico se sustenta en varios campos interdisciplinarios relacionados con el análisis de desastres naturales, la segmentación de datos y la visualización interactiva de información para la toma de decisiones.

### **Análisis de Datos en Power BI**

Power BI es una herramienta de inteligencia empresarial desarrollada por Microsoft que permite realizar análisis de datos de manera dinámica y visual. Esta plataforma integra diversas fuentes de datos, como bases de datos, archivos Excel, servicios en la nube, entre otros, y ofrece funcionalidades avanzadas para el procesamiento y visualización interactiva de la información.

Entre sus principales componentes se encuentran Power Query (para la transformación de datos), el modelo de datos (para establecer relaciones y medidas), y el entorno de visualización, que permite crear dashboard e informes personalizados. Power BI también utiliza el lenguaje DAX (Data Analysis Expressions) para cálculos y medidas complejas.

Gracias a su interfaz intuitiva, capacidades de automatización y opciones de colaboración en línea, Power BI se ha consolidado como una herramienta clave para facilitar el análisis de datos y apoyar la toma de decisiones estratégicas en tiempo real. (Microsoft, 2023)

### **Ciencias del Análisis de Datos y Minería de Datos**

El análisis de datos históricos de desastres naturales permite identificar patrones significativos en el tiempo. Esto incluye:

- Estadística Descriptiva y Exploratoria: Herramientas que permiten visualizar la frecuencia, magnitud y tendencias temporales de los desastres. (Chen, 2012)
- Minería de Datos: Métodos que identifican correlaciones ocultas entre factores como ubicación geográfica, época del año y tipo de desastre.

- **Análisis de Clústeres:** Uso de técnicas para segmentar regiones o tipos de desastres según sus características y frecuencia, permitiendo una mejor comprensión de los riesgos y la focalización de estrategias de mitigación. (Han, 2011)

### **Segmentación de Datos y Análisis de Clústeres**

El análisis de clústeres es una metodología que permite agrupar elementos con características similares, facilitando la identificación de patrones y la toma de decisiones informadas. En el contexto de desastres naturales, se utiliza para:

**Clasificación de Regiones según Nivel de Riesgo:** Segmentar zonas geográficas en función de la recurrencia y severidad de los desastres. (Cutter, 2008)

**Agrupación de Tipos de Desastres:** Identificar similitudes entre distintos eventos para entender su comportamiento y frecuencia.

**Optimización de Estrategias de mitigación:** Permitir a las autoridades focalizar recursos y diseñar políticas específicas para cada grupo de riesgo. (Levy, 2005)

### **Visualización de Datos e Inteligencia de Negocios**

La visualización de datos facilita la comprensión de patrones y tendencias complejas a través de herramientas interactivas. Aspectos relevantes incluyen:

- **Visualización geográfica:** Herramientas como Power BI y Tableau permiten superponer datos sobre mapas para identificar zonas de mayor vulnerabilidad.
- **Análisis interactivo:** Los usuarios pueden explorar datos históricos, realizar consultas personalizadas y observar la segmentación de riesgos según distintas categorías.
- **Inteligencia de negocios:** Power BI, al integrar análisis en tiempo real, apoya la toma de decisiones basadas en evidencia para diseñar políticas y medidas de mitigación más efectivas.

Este marco teórico proporciona las bases necesarias para el estudio de desastres naturales en Colombia, permitiendo identificar patrones, segmentar información crítica y mejorar la respuesta ante eventos adversos mediante el uso de tecnologías avanzadas.

### **Zonas de Amenaza y Riesgo**

Colombia se localiza dentro de una de las zonas sísmicas más activas de la Tierra, pues en la región convergen las placas tectónicas de Nazca y del Caribe contra la placa suramericana. La interacción continua entre las placas Nazca y suramericana que se mueven y chocan entre sí, ha producido la formación de montañas, cordilleras y fallas geológicas. (IDIGER, 2024)

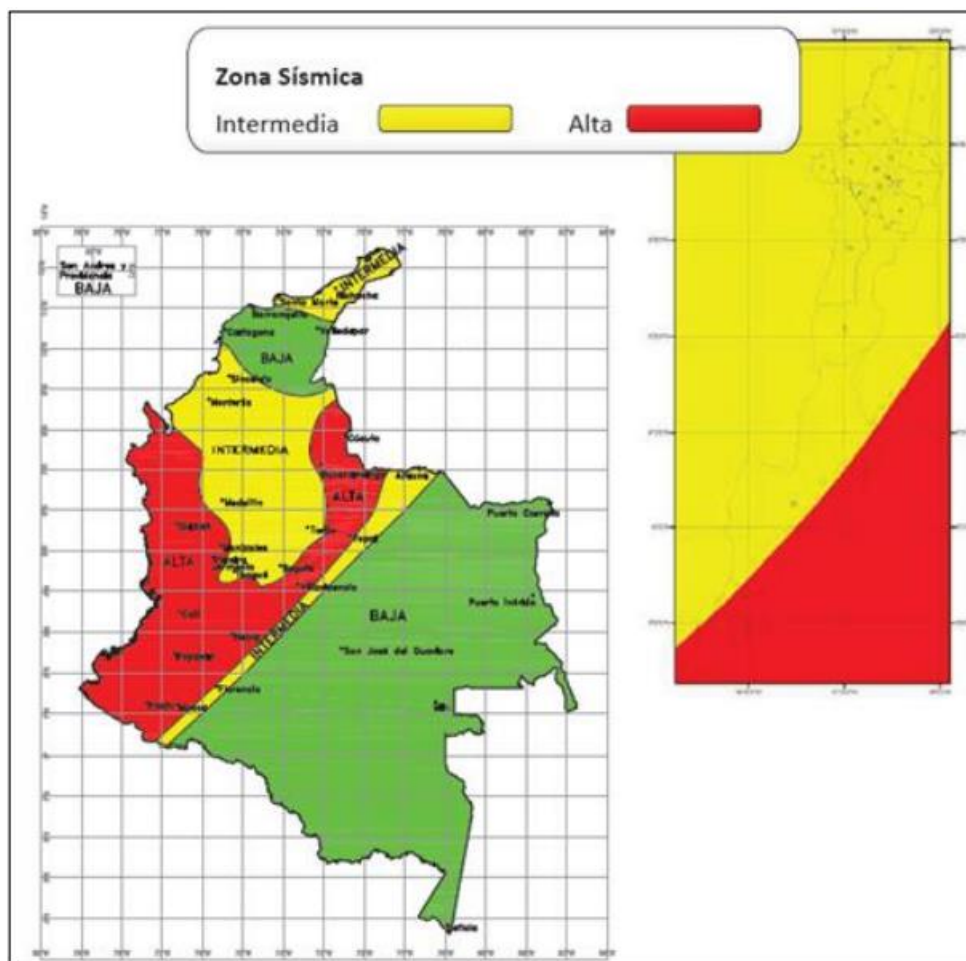
En la costa pacífica se presenta un proceso de subducción en el cual la placa Nazca se introduce debajo de la placa suramericana, el material rocoso entra de nuevo en el manto, que se encuentra a altas temperaturas, y allí se funde para luego volver a salir a través de los volcanes tales como los de la Cordillera Central del territorio colombiano. (IDIGER, 2024)

El NSR-10 (Reglamento Colombiano de Construcción Sismo Resistente) establece la clasificación del territorio colombiano en zonas según el nivel de amenaza sísmica, con el fin de definir los requisitos técnicos que deben cumplir las edificaciones para garantizar su seguridad estructural:

- Zona de amenaza sísmica baja
- Zona de amenaza sísmica intermedia
- Zona de amenaza sísmica alta

**Figura 1**

*Mapa de Zonas de Amenaza Sísmica NSR-10*



*Nota.* Tomado de Norma Sismo Resistente NSR-10

## **Metodología**

El desarrollo del proyecto seguirá un enfoque mixto (cuantitativo y cualitativo). El cual estará planteado en cuatro capítulos, que se detallaran a continuación:

### **Recolección y Preparación de Datos**

#### ***Fuentes de Datos***

Las Bases de Datos que se usaran en la ejecución de este documento, fueron obtenidas de la Pagina “Datos Abiertos Colombia”. Se encontraron cinco archivos los cuales se mostrarán en la ejecución del proyecto.

#### ***Variables para Analizar***

Tipos de desastres, ubicación, magnitud, frecuencia, impacto económico y humano.

#### ***Procesamiento Inicial***

Limpieza, estandarización y consolidación de los datos para garantizar su integridad y coherencia.

### **Consolidación de la Información y Segmentación de Datos**

#### ***Análisis Exploratorio***

Identificación de patrones históricos mediante estadísticas descriptivas y visualización de tendencias.

#### ***Técnicas Analíticas***

Aplicación de métodos de minería de datos y análisis de clústeres para segmentar las regiones y tipos de desastres según sus características y frecuencia.

#### ***Interpretación de Resultados***

Evaluación de los grupos identificados para la toma de decisiones en estrategias de mitigación.

## **Diseño y Visualización en Power BI**

### ***Herramienta Principal***

Power BI para crear paneles interactivos y visualizaciones dinámicas.

### ***Funcionalidades Clave***

Representación geoespacial de datos, análisis temporal y segmentación de desastres según características clave.

### ***Pruebas de Usabilidad***

Evaluación con planificadores de contingencias y autoridades locales para garantizar la utilidad de la plataforma.

## **Evaluación y Conclusiones**

### ***Comparación de Enfoques***

Evaluar la efectividad del análisis de clústeres y mostrar la segmentación en relación con estrategias tradicionales.

### ***Retroalimentación***

Recopilar datos y resultados obtenidos de la plataforma para posteriormente, realizar una comparación con el esquema ya establecido por la NSR-10.

### ***Documentación***

Generar un informe final que incluya los resultados, análisis de impacto y recomendaciones para su implementación, junto con los gráficos más relevantes del proyecto.

Este marco metodológico proporciona un enfoque estructurado para la recolección, análisis y visualización de datos sobre desastres naturales, permitiendo una mejor comprensión de los riesgos y facilitando la toma de decisiones basada en datos históricos obtenidos de “Datos Abiertos Colombia”.

## **Limpieza de Bases de Datos**

La preparación del conjunto de datos fue una etapa fundamental para garantizar la calidad del análisis posterior. Inicialmente, se realizó una exploración general del dataset suministrado por la Unidad Nacional para la Gestión del Riesgo de Desastres (UNGRD), el cual contenía información desde el año 2015 hasta 2022, desagregada por evento, fecha, ubicación y afectaciones.

Uno de los principales problemas detectados fue la presencia de valores nulos en variables críticas como heridos, fallecidos, desaparecidos, daños a viviendas, vías y puentes.

En este contexto, se optó por imputar el valor cero (0), interpretando que la ausencia de dato representa que no hubo afectación reportada. Esta decisión fue respaldada por la lógica del registro de emergencias, donde los ceros expresan explícitamente la no ocurrencia de daño.

Además, se detectaron múltiples inconsistencias en variables categóricas, como la denominación de los eventos. Por ejemplo, “inundación”, “Inundaciones”, “INUNDACION” y “inundacion” se trataban como registros diferentes. Para resolver este problema, se aplicó un proceso de normalización semántica, unificando todas las variantes en una categoría única con nomenclatura estandarizada. Este mismo proceso se aplicó en variables como "departamento", "municipio" y "evento".

También se eliminaron registros duplicados y se verificaron las relaciones entre fechas y ubicaciones, lo que permitió depurar los casos atípicos. Como resultado, se logró consolidar un conjunto de datos homogéneo, con más de 29.000 registros válidos listos para su análisis.

### **Limpieza de las Bases de Datos Previo a la Consolidación**

Antes de consolidar los distintos archivos de datos recopilados, se realizó un proceso sistemático de limpieza para garantizar la calidad, coherencia y compatibilidad entre las bases.

Esta fase se centró en preparar cinco archivos provenientes de la Unidad Nacional para la Gestión del Riesgo de Desastres (UNGRD), que abarcaban el periodo de 2015 a 2022. Estos archivos presentaban similitudes estructurales, aunque también diferencias importantes en la cantidad y formato de variables.

El proceso inició con la identificación y selección de variables relevantes. Se descartaron aquellas columnas que no aportaban valor al análisis o que presentaban altos niveles de valores nulos. Las variables seleccionadas fueron principalmente las relacionadas con ubicación (DIVIPOLA, departamento, municipio), fecha, tipo de evento, y afectaciones humanas y materiales (heridos, fallecidos, viviendas dañadas, entre otras).

Posteriormente, se abordó la normalización del formato de fecha, ya que los registros estaban en formato timestamp (con fecha y hora). Para el análisis, solo se consideró la fecha, por lo cual se eliminó la hora. Esto permitió simplificar el procesamiento y facilitar el análisis temporal.

Otro paso clave fue el tratamiento de valores faltantes. En lugar de eliminar registros, se optó por imputar valores cero (0) en variables cuantitativas como fallecidos o daños materiales, bajo la lógica de que la ausencia de datos indicaba que no se presentaron afectaciones en ese aspecto particular.

Finalmente, se llevó a cabo una eliminación de registros duplicados y una revisión de la consistencia semántica en variables categóricas, como “evento”. Esto incluyó corregir errores ortográficos y unificar nomenclaturas para asegurar que categorías equivalentes no se trataran como diferentes.

Todo este proceso sentó las bases para la consolidación de un único conjunto de datos limpio, uniforme y confiable, que sería utilizado en el desarrollo de las visualizaciones y los análisis posteriores en Power BI

Finalmente dejando así la base de datos consolidada con las siguientes variables:

DIVIPOLA “División Político-Administrativa de Colombia”	Desaparecidos
Fecha	Viviendas Destruidas
Estación	Viviendas Averiadadas
Departamento	Vías
Municipio	Puentes
Evento	Centros de Salud
Fallecidos	Centros Educativos
Heridos	Centros Comunitarios
	Hectáreas Afectadas

## Visualización en Power BI

La visualización de datos constituyó una fase clave del proyecto, ya que permitió representar gráficamente la información consolidada sobre desastres naturales ocurridos en Colombia entre 2015 y 2022. Para ello, se implementó una plataforma interactiva utilizando Power BI, con el propósito de facilitar el análisis, la exploración dinámica y la comunicación efectiva de los hallazgos.

El tablero fue estructurado en múltiples hojas temáticas que incluyen visualizaciones como mapas geoespaciales, gráficos de barras, diagramas circulares, líneas temporales y paneles de resumen. Estas visualizaciones permitieron observar la distribución de los eventos por tipo, frecuencia, ubicación geográfica, y su evolución a lo largo del tiempo. Se incorporaron filtros por año, trimestre, tipo de amenaza y departamento, lo que ofreció al usuario la posibilidad de segmentar los datos según diferentes criterios analíticos, como se visualiza en la *Figura 2*.

Entre los elementos destacados se encuentran: indicadores de eventos más frecuentes, departamentos más afectados, y la evolución mensual y trimestral de los desastres. Adicionalmente, se integraron algoritmos de agrupamiento mediante el uso de scripts de Python (K-Modes), lo cual permitió segmentar el territorio en clústeres con características similares de exposición y recurrencia.

**Figura 2**

*Visualización Resumen Geográfico de Eventos*



*Nota.* Visualización resumen geográfico de eventos: ubicación, total de registros, evento más frecuente y departamento más afectado.

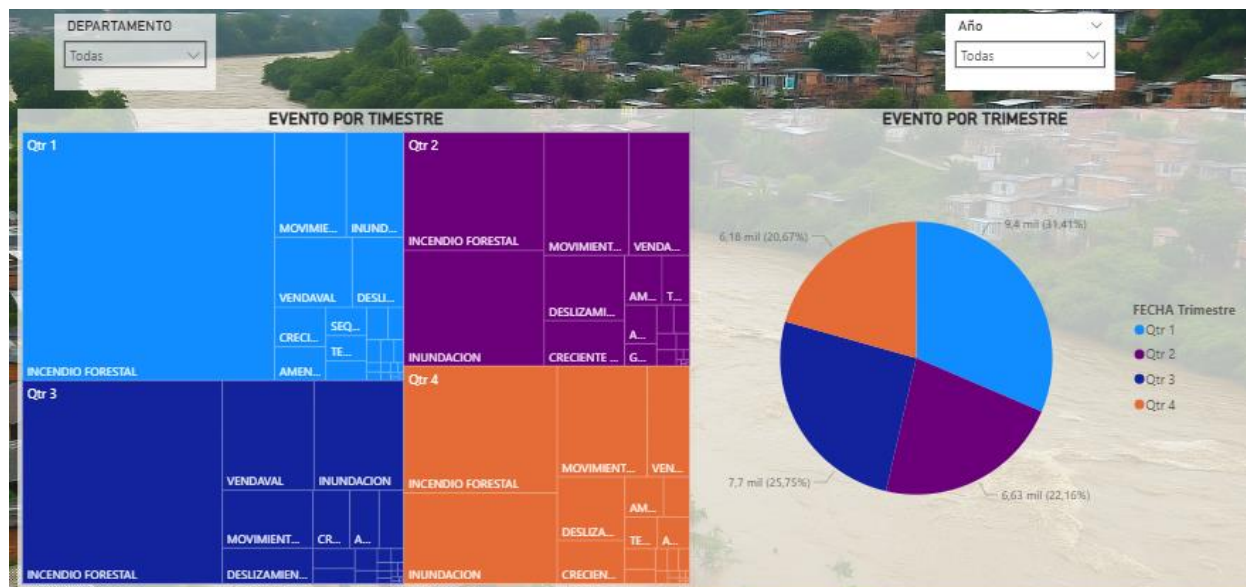
La utilización de funciones DAX facilitó la creación de medidas dinámicas, como el evento más recurrente o el total de registros por categoría, mejorando la interactividad del tablero, como se puede observar en la *Figura 3* Esta fase de visualización no solo enriqueció el análisis descriptivo, sino que también proporcionó una herramienta poderosa para la toma de decisiones en materia de gestión del riesgo, gracias a su claridad visual, flexibilidad de exploración y capacidad de síntesis.

La visualización de datos es fundamental para interpretar y comunicar hallazgos de manera efectiva. Este segmento del proyecto se dedicará a la utilización de Power BI como herramienta principal para crear visualizaciones interactivas que faciliten la exploración de los

datos. Se discutirán las características de Power BI que permiten una representación visual dinámica y se presentarán ejemplos de visualizaciones desarrolladas a partir de la base de datos consolidada.

**Figura 3**

*Visualización Eventos por Trimestre: Distribución y Proporción Según Tipo de Desastre*



## **Análisis de Datos y Graficas**

El análisis de datos y visualizaciones constituye un componente fundamental de este proyecto, orientado a extraer conclusiones significativas a partir de la información representada gráficamente. En esta sección se abordarán tanto las técnicas de análisis empleadas como los hallazgos más relevantes obtenidos a través de las gráficas generadas en Power BI. Se hará especial énfasis en la interpretación de los resultados, destacando su utilidad dentro del contexto de la gestión del riesgo de desastres naturales. Esta etapa del estudio no solo permite identificar patrones y tendencias, sino que también provee herramientas clave para la toma de decisiones informadas en la planificación y prevención de eventos adversos.

### **Análisis Descriptivo**

A partir de la información consolidada, se desarrolló una visualización geoespacial en forma de mapa, donde el tamaño de las burbujas representa la concentración de desastres naturales registrados entre los años 2015 y 2022. Esta representación permite identificar, de manera clara, que Cundinamarca es el departamento con el mayor número de eventos reportados durante el período analizado.

Para enriquecer el análisis, se integraron tres tarjetas informativas clave que facilitan una lectura más detallada de los datos: la primera refleja el total acumulado de eventos, que alcanza los 29.913 casos; la segunda señala que el tipo de desastre más frecuente es el incendio forestal; y la tercera corrobora que Cundinamarca es, efectivamente, el departamento más afectado en términos de frecuencia de desastres naturales.

**Figura 4**

*Análisis De Datos Predominantes En Base De Datos*



Gracias a las funcionalidades que ofrece la plataforma Power BI, es posible desglosar y explorar en detalle los datos mencionados previamente. Esto permite visualizar la información de manera dinámica, segmentándola por cada año del periodo de estudio. Como resultado, se consolida la siguiente tabla comparativa que resume los principales indicadores anuales.

**Tabla 1**

*Consolidado Comparativo de los Indicadores Anuales*

Año	Total, Eventos	Evento Frecuente	Dpto. más afectado
2015	3.528	Incendio Forestal	Cundinamarca
2016	3.736	Incendio Forestal	Cundinamarca
2017	3.182	Incendio Forestal	Cundinamarca

Año	Total, Eventos	Evento Frecuente	Dpto. más afectado
2018	3.343	Incendio Forestal	Cundinamarca
2019	4.141	Incendio Forestal	Cundinamarca
2020	3.642	Incendio Forestal	Antioquia
2021	3.877	Movimiento en masa	Cundinamarca
2022	4.464	Movimiento en masa	Cundinamarca

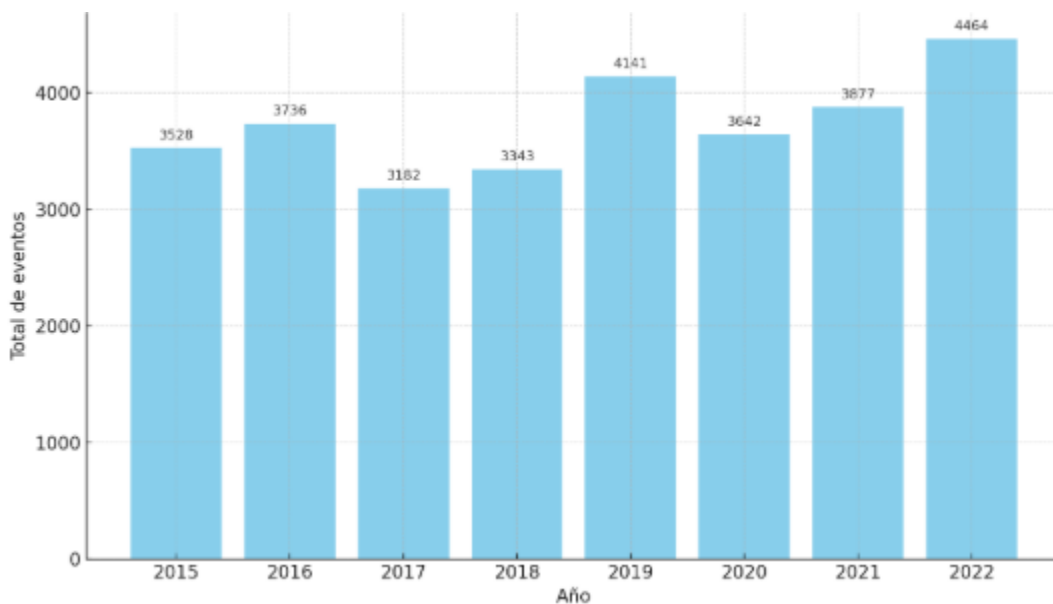
A partir de la tabla anterior, se puede observar que, a lo largo de los años, los *incendios forestales* y los *movimientos en masa* (como deslizamientos) se han mantenido como amenazas constantes en el territorio colombiano. Este patrón evidencia la necesidad de priorizar estrategias de prevención y mitigación frente a estos tipos de eventos. Asimismo, los departamentos de *Cundinamarca* y *Antioquia* se destacan de forma recurrente como los más afectados, lo que los convierte en zonas estratégicas para focalizar *planes de gestión del riesgo y respuesta ante desastres*.

Adicionalmente de acuerdo con la *Tabla 1*., podemos generar una gráfica de barra como se visualiza en la *Figura 5*, donde se visualiza que, a lo largo del periodo analizado, se presenta una tendencia general creciente en la cantidad total de desastres reportados en Colombia, alcanzando su punto más alto en el año 2019, con más de 4.141 eventos registrados. Este incremento podría reflejar tanto un aumento real en la ocurrencia de desastres como una mejora en los sistemas de reporte y monitoreo. Sin embargo, en 2020 se presenta una disminución significativa en el número de eventos, lo cual podría estar relacionado con los efectos de la pandemia por COVID-19 que impactó las dinámicas de reporte y atención institucional. A partir de 2021, el volumen de eventos vuelve a incrementarse, lo que podría sugerir un retorno a la

normalidad en la gestión de estos datos o un aumento en la severidad y frecuencia de los fenómenos naturales.

### Figura 5

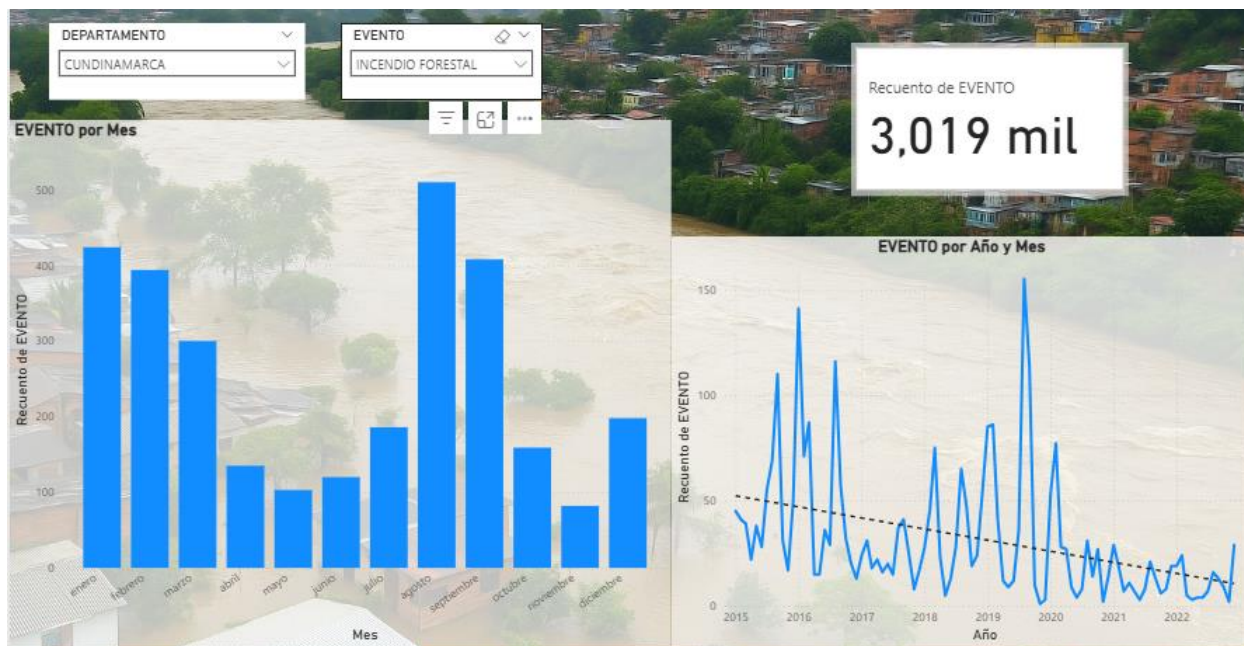
*Total de Eventos por Año*



De acuerdo con lo analizado anteriormente y observando que los departamentos de *Cundinamarca* y *Antioquia* se destacan de forma recurrente como los más afectados y apoyándonos con las *Tabla 1*, filtramos las visualizaciones por departamentos y evento donde se obtiene la *Figura 6* y *Figura 7*.

**Figura 6**

*Eventos por Año y Mes para Cundinamarca vs Incendio Forestal*



En la visualización de la *Figura 6*, corresponde al departamento de Cundinamarca y por el desastre natural más común identificado por los pasos anteriores, observamos en el gráfico de barras, que resume el recuento de eventos por mes, donde se evidencia que en los meses donde se presentan los picos de más altos de presencia del desastre natural son enero-febrero y agosto-septiembre lo cual esto nos permitirá ajustar y ayudar a optimizar el plan de contingencia frente a un desastre natural.

El gráfico de líneas, que representa el número de eventos por año y mes, complementa esta visión mostrando una alta estacionalidad. Se identifican picos mensuales pronunciados, especialmente entre 2016 y 2020, con algunos meses que superan los 150 eventos. Estos picos podrían corresponder a temporadas climáticas críticas, como las épocas de lluvia más intensas en

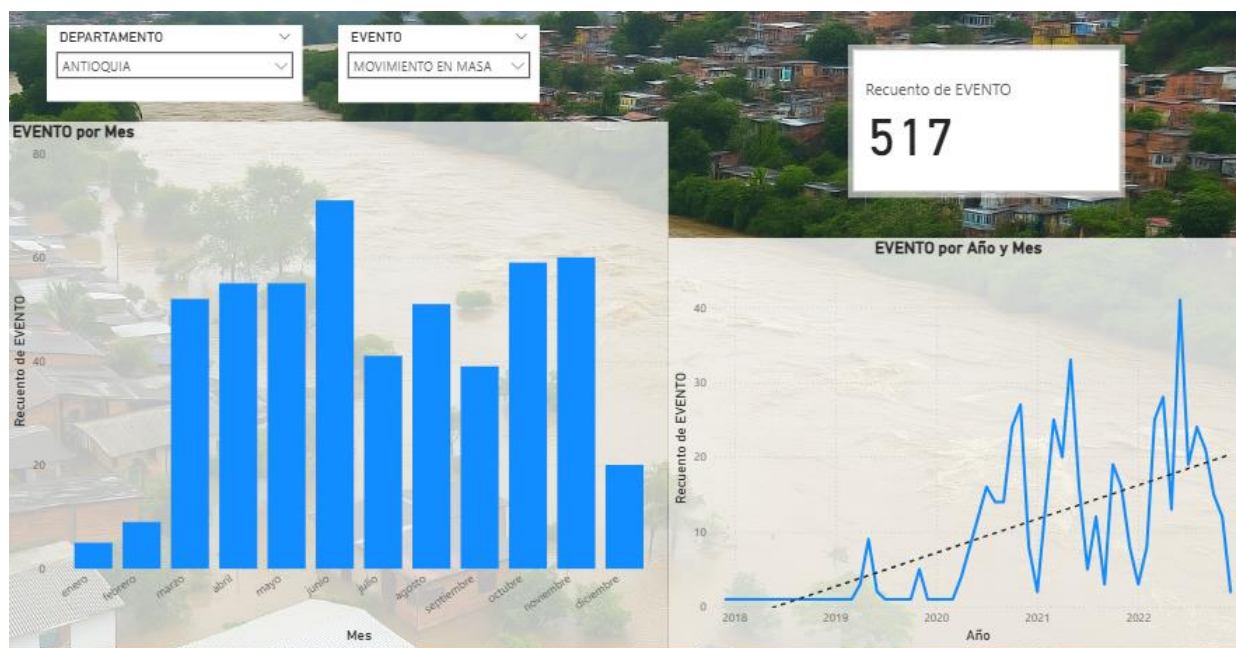
la región. A partir de 2021, los picos disminuyen en intensidad, indicando una posible estabilización o menor severidad de los eventos registrados. También se muestra una línea punteada que sugiere un promedio mensual cercano a los 50 eventos, lo cual puede servir como referencia para detectar desviaciones significativas, adicionalmente se evidencia que 2016 fue el año con mayor número de eventos, superando los 850 registros. A lo largo del periodo observado, los datos muestran una tendencia oscilante, con variaciones significativas entre un año y otro. Por ejemplo, se observa una caída importante en el año 2020, lo cual podría estar relacionado con el impacto de la pandemia por COVID-19, que afectó no solo la ocurrencia sino también el reporte de eventos en todo el país. Posteriormente, en 2021 y 2022, se nota una recuperación en el número de registros, aunque sin alcanzar los niveles máximos de años anteriores.

En la parte superior derecha del tablero se encuentra un contador que resume el total de eventos registrados durante el periodo analizado, con una cifra de 3.019 eventos (Incendio Forestal) y para este caso en el departamento de Cundinamarca.

En resumen, el análisis evidencia un comportamiento variable en la ocurrencia de eventos en Cundinamarca durante los últimos ocho años, con un año pico en 2016 y una caída notable en 2020, seguida de una recuperación parcial. Además, se destacan patrones estacionales donde se presenta los picos más altos en enero-febrero y agosto-septiembre que pueden estar relacionados con factores climáticos recurrentes. Este tipo de análisis es útil para la planificación preventiva, asignación de recursos y diseño de políticas públicas orientadas a la gestión del riesgo en el territorio.

**Figura 7**

*Eventos por Año y Mes para Antioquia vs Movimiento en Masa*



En la visualización de *Figura 7*, corresponde al departamento de Antioquia y por el desastre natural más común identificado por los pasos anteriores, se evidencia una marcada tendencia al alza en la ocurrencia de desastres naturales durante el periodo comprendido entre 2015 y 2022. El gráfico de barras muestra que, que lo meses donde se presenta el pico más alto de la presencia del desastre natural son en los meses de junio, octubre y noviembre donde junio es el mes donde se obtuvo la mayor presencia de dicho desastre natural, por otra parte, se puede observar que la frecuencia de eventos se mantuvo relativamente estable, con ligeras variaciones anuales, a partir del año 2020 se presenta un incremento significativo en la cantidad de eventos registrados. Este crecimiento continúa durante 2021 y alcanza su punto máximo en 2022, año en el que se contabilizan más de 500 eventos, lo que indica una intensificación notable de los fenómenos naturales o un mejor sistema de reporte en el departamento.

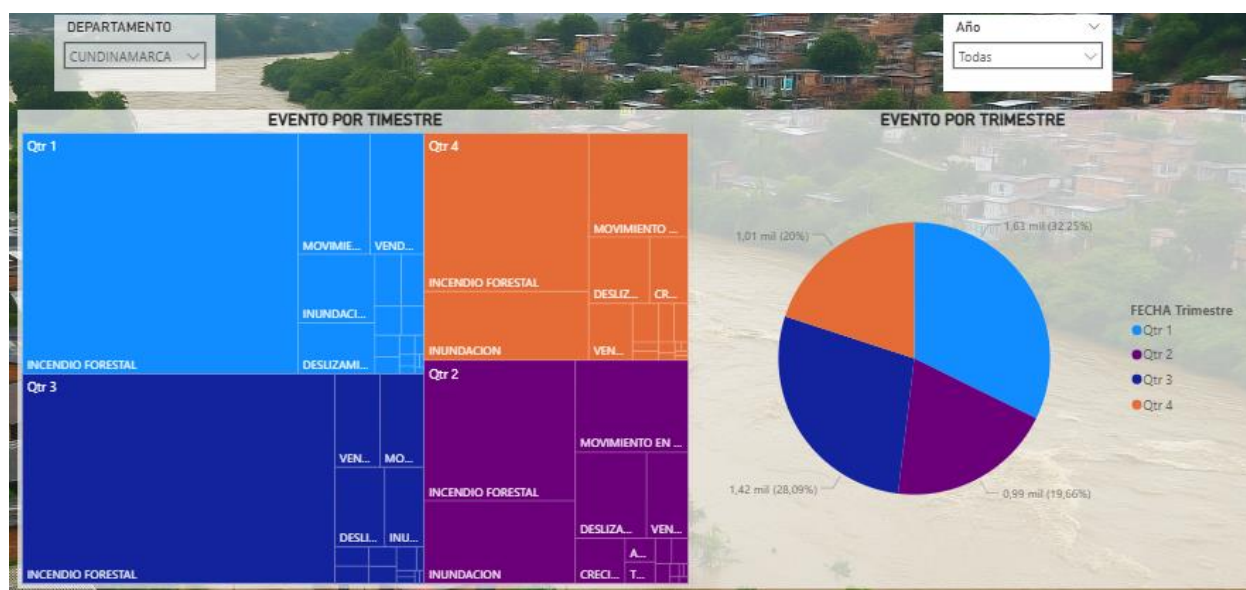
Por su parte, el gráfico de líneas que representa los eventos por mes y año refuerza esta tendencia creciente, mostrando además una mayor variabilidad mensual en los años más recientes.

Desde 2020 se observan picos pronunciados que en algunos meses superan los 80 eventos, lo cual sugiere una posible relación con eventos climáticos extremos como el fenómeno de La Niña. Este comportamiento no solo representa una carga considerable para los sistemas de atención y mitigación de emergencias, sino que también pone en evidencia la necesidad de estrategias más robustas y adaptativas para la gestión del riesgo.

Por otra parte, mediante la en Figura 8, el departamento de *Cundinamarca* se ha consolidado como el más afectado por desastres naturales a nivel nacional. A través del análisis visual de datos anuales por trimestres, se evidencia una alta concentración de eventos, siendo los más recurrentes los *incendios forestales*, seguidos por inundaciones, vendavales y deslizamientos.

## Figura 8

*Evento por Trimestre para Cundinamarca*



Al observar la distribución trimestral de estos eventos, se identifican patrones estacionales claros. El *primer trimestre (Q1)*, correspondiente a los meses de enero a marzo, muestra una alta prevalencia de incendios forestales, situación atribuida a condiciones climáticas secas y la acumulación de material vegetal combustible tras la temporada de lluvias. En varios años, este trimestre también registra vendavales y algunos episodios de heladas.

Durante el *segundo trimestre (Q2)*, que abarca de abril a junio, se presenta una transición hacia eventos asociados a lluvias intensas. Aquí se incrementan los *deslizamientos, inundaciones y crecientes súbitas*, en respuesta a la temporada de lluvias en la región andina. Aunque los incendios forestales aún están presentes, su frecuencia disminuye frente a las amenazas hidrometeorológicas.

El *tercer trimestre (Q3)*, entre julio y septiembre, representa uno de los períodos de mayor criticidad. Se destaca nuevamente el aumento de *incendios forestales*, con picos notables como en el año 2019, donde se registraron más de 300 incendios solo en este trimestre. Además, persisten vendavales e inundaciones en algunas zonas, mostrando una coexistencia de amenazas climáticas.

Por su parte, el *cuarto trimestre (Q4)*, de octubre a diciembre, muestra una *mayor diversidad en los tipos de eventos*. Se observan incendios forestales, movimientos en masa, crecientes súbitos y eventos menos frecuentes como granizadas y tormentas. A partir del año 2020, se nota un incremento en los “movimientos en masa” y otras amenazas asociadas a la saturación de suelos tras largos períodos de precipitación.

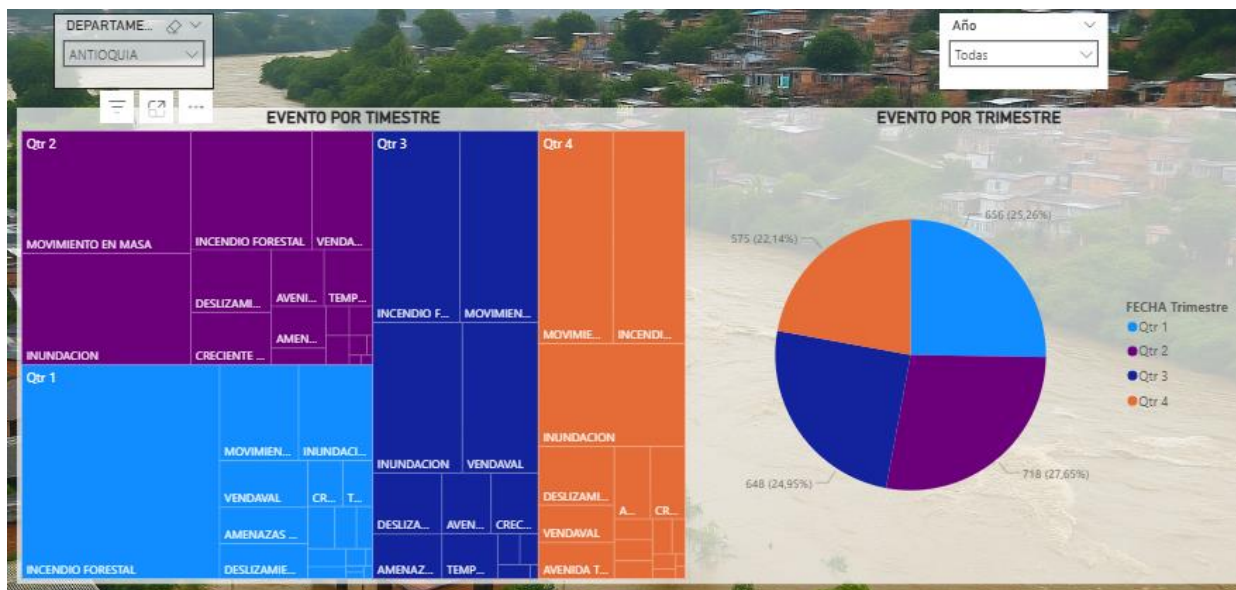
En conclusión, el comportamiento de los desastres naturales en Cundinamarca presenta una clara *estacionalidad*, con eventos predominantes que varían según el trimestre. La frecuencia

sostenida de incendios forestales y la emergencia de nuevos tipos de amenazas en los últimos años hacen evidente la necesidad de fortalecer los sistemas de alerta temprana, las estrategias de gestión del riesgo y las campañas de prevención comunitaria, particularmente en los trimestres más críticos como el primero y el tercero.

Por otra parte, frente a la visualización de la *Figura 9*, podemos analizar durante el período analizado, el departamento de *Antioquia* ha enfrentado una notable cantidad de desastres naturales, con una distribución trimestral bien diferenciada. La gráfica de treemap combinada con el gráfico de pastel evidencia una alta diversidad de amenazas, siendo las más relevantes las *inundaciones*, *incendios forestales*, *movimientos en masa* y *vendavales*. En total, se registraron aproximadamente *2.597 eventos*, con una distribución bastante equilibrada a lo largo del año.

## Figura 9

### Evento por Trimestre para Antioquia



El primer trimestre (Q1) representa el 25,26% del total de eventos (655 casos), destacándose los *incendios forestales* como la amenaza predominante, lo cual es coherente con el

comportamiento climático seco habitual de inicios de año. También son relevantes los *vendavales, deslizamientos e inundaciones*, lo que indica una mezcla de amenazas climáticas tanto secas como húmedas en este período.

En el *segundo trimestre (Q2)*, con el *22,14% de los eventos (575 casos)*, se observa una transición clara hacia eventos típicos de la temporada de lluvias. Aquí se intensifican *los movimientos en masa, las inundaciones y las crecientes súbitas*, lo que responde a la saturación de los suelos y al incremento del caudal de ríos. Los *incendios forestales* aún tienen presencia, pero en menor proporción que en el primer trimestre.

El *tercer trimestre (Q3)* agrupa el *24,96% de los eventos (648 casos)*. Este período es crítico para Antioquia, con una presencia simultánea de múltiples amenazas. Se mantienen altas las *inundaciones* y los *movimientos en masa*, acompañados por *vendavales* e incluso tormentas eléctricas, lo que sugiere condiciones atmosféricas inestables con intensas precipitaciones.

Por su parte, el *cuarto trimestre (Q4)* concentra el *27,65% de los eventos (718 casos)*, siendo el más impactado del año. Las *inundaciones* y los *movimientos en masa* son particularmente dominantes, lo que señala una persistencia de lluvias en la región hasta finalizar el año. Además, se reportan incendios forestales, vendavales y otros eventos como avenidas torrenciales y crecientes súbitas.

En resumen, Antioquia experimenta una distribución casi uniforme de desastres naturales a lo largo de los cuatro trimestres, con picos importantes en el último trimestre del año. La constante aparición de inundaciones y movimientos en masa en casi todos los trimestres refleja una alta exposición a amenazas hidrometeorológicas. Esta situación exige una planificación integral del riesgo, con énfasis en gestión de cuencas, reforestación, control del uso del suelo y

fortalecimiento de sistemas de monitoreo en tiempo real, especialmente en las temporadas lluviosas del segundo y cuarto trimestre.

## Impactos y Afectaciones

### *Heridos, Fallecidos, Desaparecidos*

Para realizar el respectivo análisis de los heridos, fallecidos y desapariciones se presenta un resumen general del número total de heridos, fallecidos y desaparecidos por tipo de evento. Donde se evidencia que las "amenazas concatenadas o complejas" son el tipo de evento que ha generado mayor número de heridos, con un total cercano a los 1.800 casos. Le siguen eventos como incendios forestales, movimientos en masa y avenidas torrenciales, lo que sugiere que estos eventos tienen una alta capacidad de afectar la integridad física de las personas, aunque no necesariamente estén entre los más letales.

### Figura 10

#### *Resumen General del Número Total de Heridos, Fallecidos y Desaparecidos por Tipo de Evento*



De acuerdo con la *Figura 10*, respecto a los fallecimientos, el evento más letal ha sido el deslizamiento, con 641 fallecidos, seguido por avenidas torrenciales con una cantidad de 515 fallecidos. Esto refleja la peligrosidad de los eventos geológicos e hidrometeorológicos en zonas vulnerables, particularmente en contextos de lluvias intensas.

En cuanto a los desaparecidos, el mayor número de casos se registra también en el contexto de avenidas torrenciales, con más de 100 personas desaparecidas, lo cual indica no solo la severidad del fenómeno, sino también dificultades en la respuesta inmediata o rescate. Este dato es clave para las estrategias de gestión del riesgo y respuesta humanitaria.

Gracias a las funcionalidades que ofrece la plataforma Power BI, es posible desglosar y explorar en detalle los datos mencionados previamente. Esto permite visualizar la información de manera dinámica, segmentándola por cada año del periodo de estudio. Como resultado, se consolida la siguiente tabla comparativa que resume por año el número total de heridos, fallecidos y desaparecidos por tipo de evento.

**Tabla 2**

*Consolidado por Año el Número Total de Heridos, Fallecidos y Desaparecidos por Tipo de Evento*

Año	Desaparecidos	Fallecidos	Heridos
2015	Amenazas Concatenada	Avenida Torrencial	Incendio Forestal
2016	Amenazas Concatenada	Deslizamientos	Amenazas Concatenada
2017	Avenida Torrencial	Avenida Torrencial	Avenida Torrencial
2018	Granizada	Deslizamientos	Granizada
2019	Inundación	Movimientos En Masa	Erosión

Año	Desaparecidos	Fallecidos	Heridos
2020	Avenida Torrencial	Deslizamientos	Incendio Forestal
2021	Inundación	Movimientos En Masa	Movimientos En Masa
2022	Creciente Súbita	Movimientos En Masa	Amenazas Concatenada

La anterior tabla nos permite identificar tendencias anuales en los tipos de eventos más críticos en cada categoría. Se observa que, de forma recurrente, las avenidas torrenciales han sido los eventos que más han generado tanto desaparecidos como fallecidos, especialmente en los años 2015, 2017, y 2020. Por su parte, los movimientos en masa dominan en términos de mortalidad en años como 2016, 2018, 2019 y 2021, mostrando una presencia constante en el tiempo.

En cuanto a los heridos, se destaca que los incendios forestales han tenido un gran impacto en años como 2015 y 2020, mientras que las amenazas concatenadas o complejas aparecen repetidamente desde 2016 hasta 2022 como uno de los fenómenos que más afectan la salud y seguridad de las personas, posiblemente por su carácter multifactorial o simultáneo.

En conjunto, este análisis evidencia que ciertos eventos como las avenidas torrenciales y los movimientos en masa representan los mayores riesgos en términos de pérdida de vidas humanas y desapariciones, mientras que las amenazas complejas e incendios forestales son más relevantes en términos de lesiones. Esta diferenciación es clave para priorizar acciones de mitigación, preparación comunitaria y asignación de recursos en gestión del riesgo de desastres.

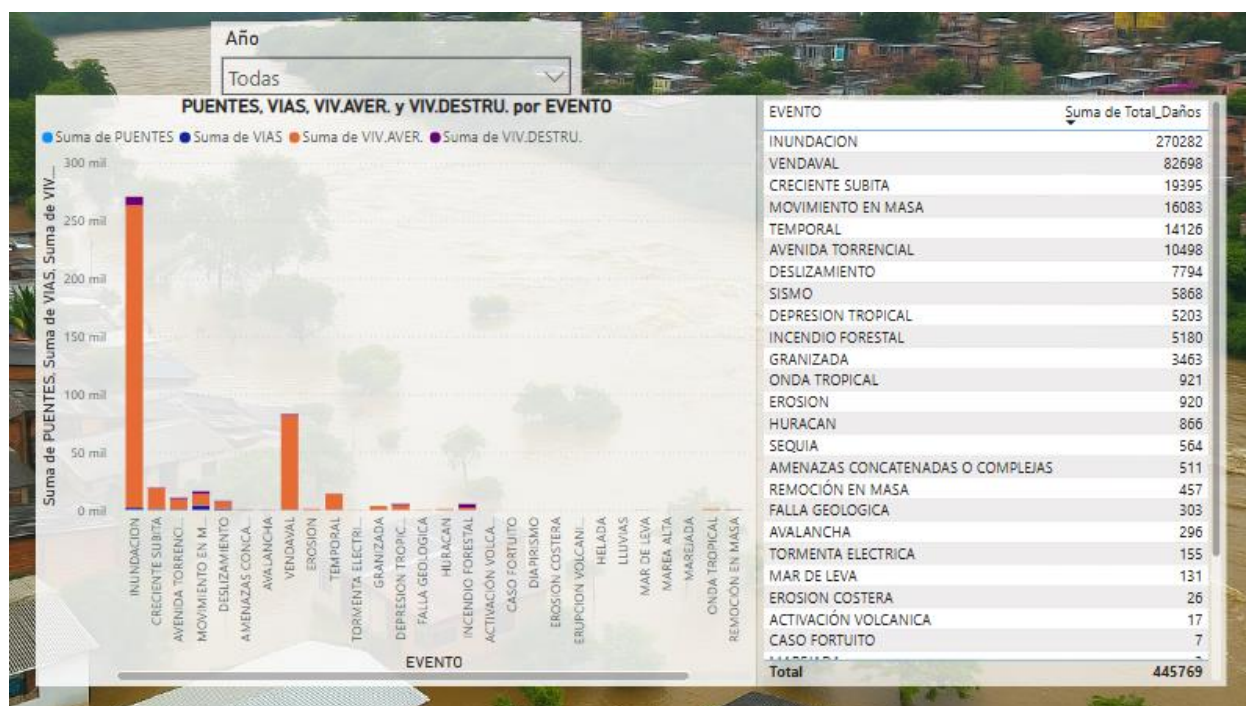
#### ***Daños a Infraestructura (Puentes, Vías, Viviendas)***

Con el propósito de identificar los desastres naturales que han tenido mayor incidencia en los daños a la infraestructura departamental, se realizó un análisis detallado de los eventos

ocurridos durante los años del periodo de estudio. La evaluación se centró en los efectos sobre elementos clave como puentes, vías y viviendas, los cuales representan componentes fundamentales para la conectividad y el bienestar de las comunidades. En este sentido, la figura presentada permite observar con mayor precisión cuáles tipos de desastres han generado los mayores niveles de afectación, brindando una perspectiva integral sobre los impactos diferenciados que cada amenaza natural produce en el territorio.

**Figura 11**

*Resumen Daños en Puentes, Vías y Viviendas*



De acuerdo con la *Figura 11*, podemos analizar que los desastres naturales han generado un impacto considerable en la infraestructura del territorio colombiano, especialmente en viviendas. A lo largo del periodo comprendido entre 2015 y 2022, las *inundaciones* se posicionan como el evento más destructivo, registrando el mayor número de daños totales

superiores a las 270 mil afectaciones, lo cual evidencia su frecuencia y severidad, y las convierte en una prioridad para la gestión del riesgo y la planificación territorial.

Gracias a las funcionalidades que ofrece la plataforma Power BI, es posible desglosar y explorar en detalle los datos mencionados previamente. Esto permite visualizar la información de manera dinámica, segmentándola por cada año del periodo de estudio. Como resultado, se consolida la siguiente tabla comparativa que resume por año total de datos en la infraestructura por tipo de evento.

**Tabla 3**

*Consolidado Años por Evento con Mayo Total de Daños en Infraestructura*

Año	Evento Mayor Daños
2015	Vendaval-Inundación
2016	Inundación-Vendaval
2017	Inundación
2018	Inundación
2019	Inundación-Vendaval
2020	Inundación
2021	Inundación
2022	Inundación

De acuerdo con la tabla anterior, se puede concluir que el desastre natural que genera la mayor afectación en la infraestructura del territorio nacional es la inundación, evidenciada por su alta incidencia en los daños registrados. No obstante, los vendavales también representan un evento de significativa recurrencia y severidad en términos de impacto estructural. Esta

información resulta clave para orientar la priorización de acciones en materia de mitigación, fortalecimiento de la preparación comunitaria y una adecuada asignación de recursos dentro de las estrategias de gestión del riesgo de desastres.

## Análisis de Clúster

El análisis de clústeres realizado en este proyecto no se limita únicamente a técnicas estadísticas descriptivas, sino que incorpora herramientas avanzadas de aprendizaje automático no supervisado, específicamente mediante el uso del algoritmo K-Modes. Este tipo de modelos permite identificar patrones ocultos en grandes volúmenes de datos categóricos, agrupando observaciones que comparten similitudes estructurales, espaciales y temáticas en torno a los desastres naturales ocurridos en Colombia.

Para llevar a cabo este análisis, se implementó un script en Python integrado a través de la función “Ejecutar script de Python” en Power BI. El modelo aplicado fue K-Modes, un algoritmo de clustering diseñado para datos categóricos, que funciona de manera similar al K-Means, pero utilizando la distancia de coincidencia simple (matching dissimilarity) en lugar de la distancia euclidiana. Esta técnica busca minimizar la disimilitud dentro de cada grupo y maximizar la diferencia entre clústeres.

El proceso de formación de los clústeres implicó:

- Seleccionar las variables categóricas más representativas: tipo de evento, departamento, trimestre y afectaciones.
- Determinar el número óptimo de clústeres mediante un análisis de costo (cost function) iterativo.
- Ejecutar el modelo con los datos históricos del periodo 2015–2022.
- Integrar los resultados al modelo Power BI para su visualización y análisis geoespacial.

Los datos fueron segmentados en cinco clústeres principales, cada uno con características únicas en términos de tipo de desastre predominante, la ubicación geográfica de los eventos, la

frecuencia de ocurrencia, el nivel de impacto potencial y la complejidad asociada a la gestión del riesgo. Esta segmentación facilita la toma de decisiones orientadas a la prevención, asignación de recursos y planeación territorial diferenciada.

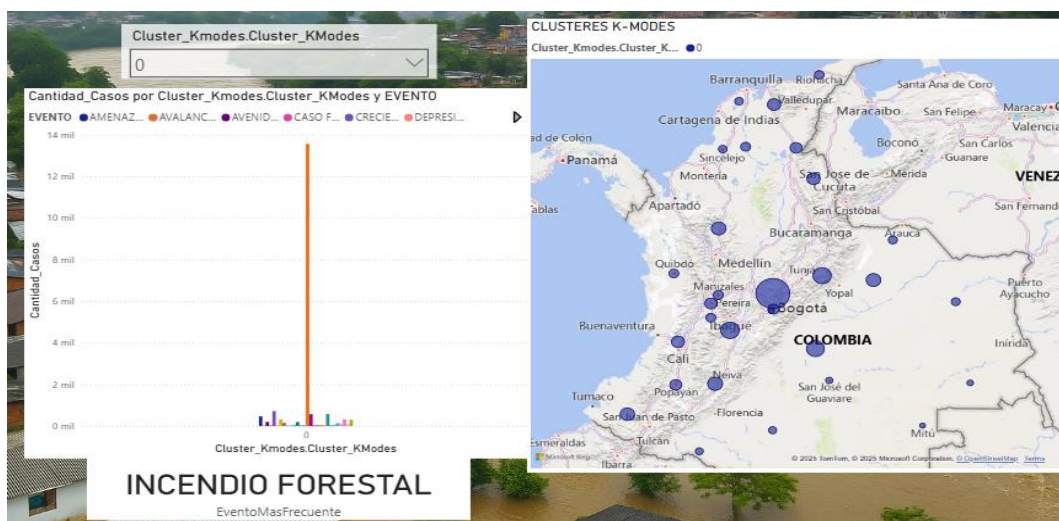
## Características de Cada Grupo

### *Clúster 0 – Incendios Forestales como Evento Dominante*

El Clúster 0 concentra el mayor volumen de eventos naturales en el conjunto de datos, con un total de 17.896 casos. Al analizar el gráfico de barras, se observa un predominio abrumador del evento “incendio forestal”, el cual representa más del 70% del total del clúster. Este tipo de evento muestra una recurrencia estacional clara, asociada a los trimestres secos (Q1 y Q3), en donde las condiciones climáticas como altas temperaturas, baja humedad y fuertes vientos incrementan la probabilidad de ignición. La uniformidad de eventos dentro del clúster evidencia un perfil de amenaza bien definido, lo que permite caracterizarlo como una categoría de riesgo climático estacional homogéneo.

## Figura 12

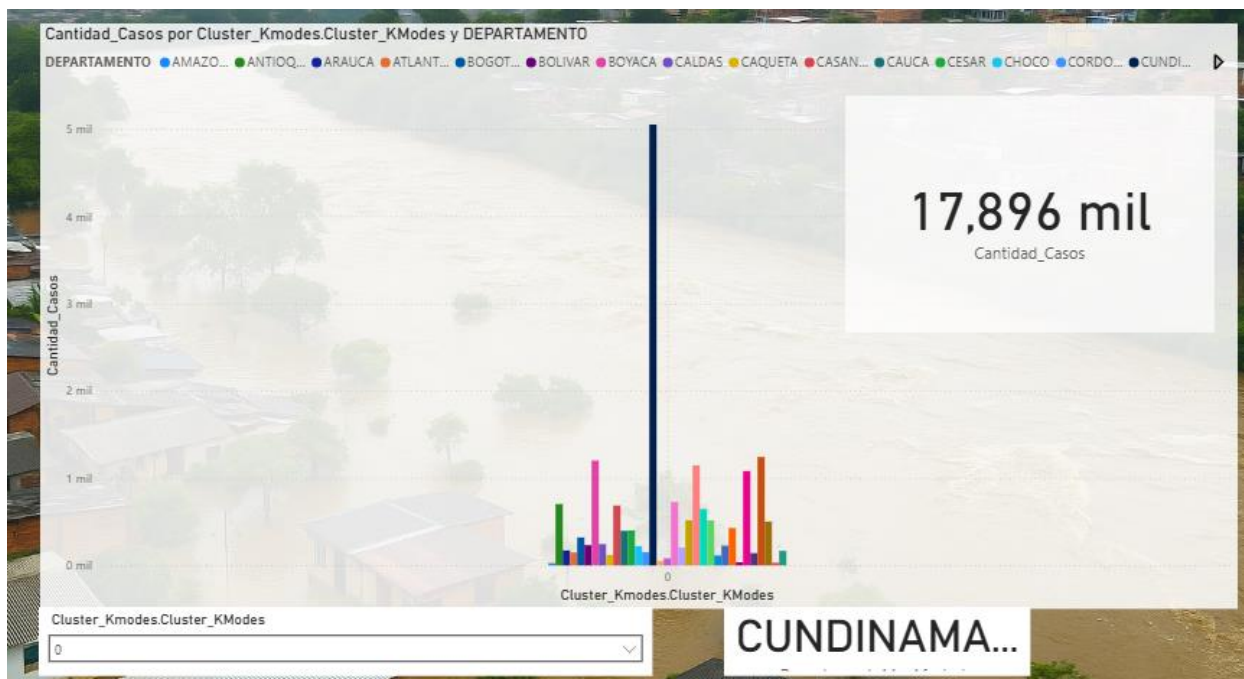
### *Visualización de Clúster 0*



Desde el punto de vista geográfico, el mapa de burbujas y como lo ratifica la *Figura 13*, revela que estos eventos están altamente concentrados en la región andina central, principalmente en *Cundinamarca, Tolima, Meta, Antioquia* y parte del *Huila*. Estas zonas, que combinan áreas rurales con alta densidad forestal y regiones periurbanas en expansión, presentan una exposición significativa a incendios forestales inducidos por actividades humanas y condiciones climáticas extremas. La identificación de estos focos territoriales permite orientar acciones preventivas como campañas educativas, restricciones de quemas controladas y fortalecimiento de cuerpos de bomberos forestales, especialmente en los trimestres donde se dispara la incidencia.

### Figura 13

#### Detalle de Clúster 0

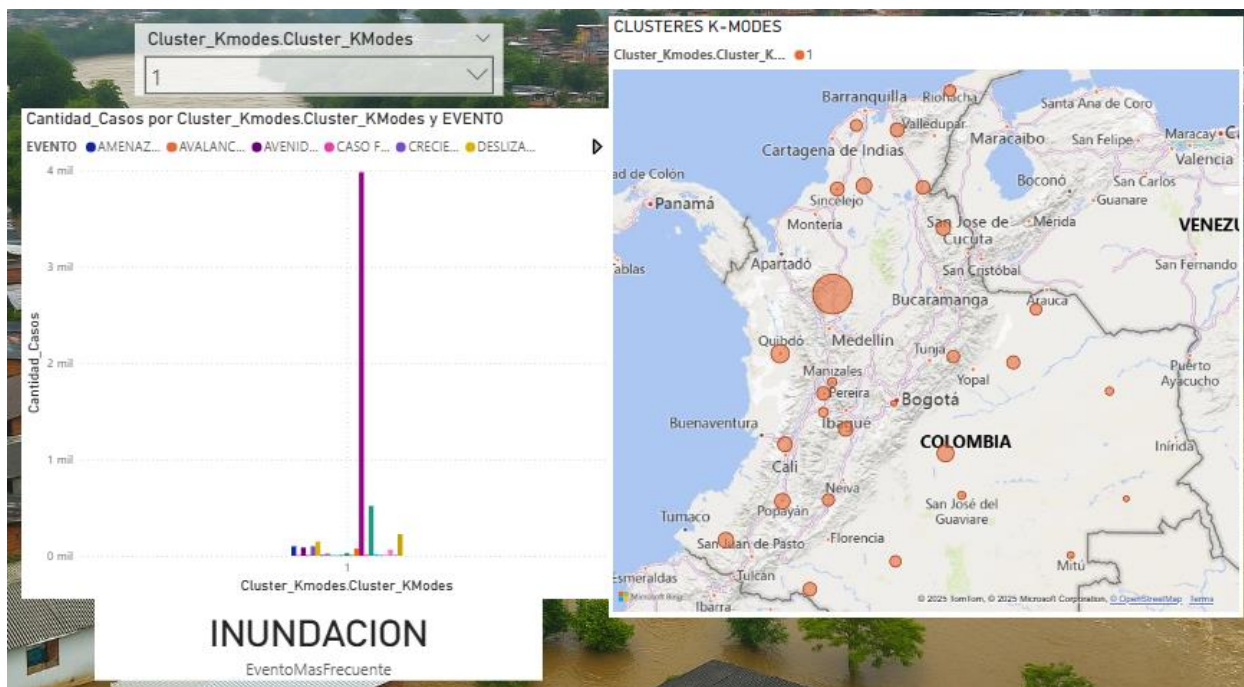


### Clúster 1–Inundaciones y Movimientos en Masa en Antioquia

Este grupo reúne eventos de carácter *hidrometeorológico*, como *inundaciones*, *crecientes súbitas* y *vendavales*, los cuales se activan principalmente en los trimestres Q2 y Q4, coincidiendo con las temporadas de lluvias más intensas en el país. En la *Figura 14*, se visualiza una gráfica de barras donde se presenta una distribución más equitativa entre los tipos de eventos, aunque las inundaciones son el fenómeno más representativo. Este clúster refleja un patrón claro de vulnerabilidad a la *dinámica hídrica* y pone de relieve la importancia de gestionar el riesgo en contextos donde el agua representa tanto una fuente de vida como una amenaza cíclica.

#### Figura 14

##### Visualización de Clúster 1

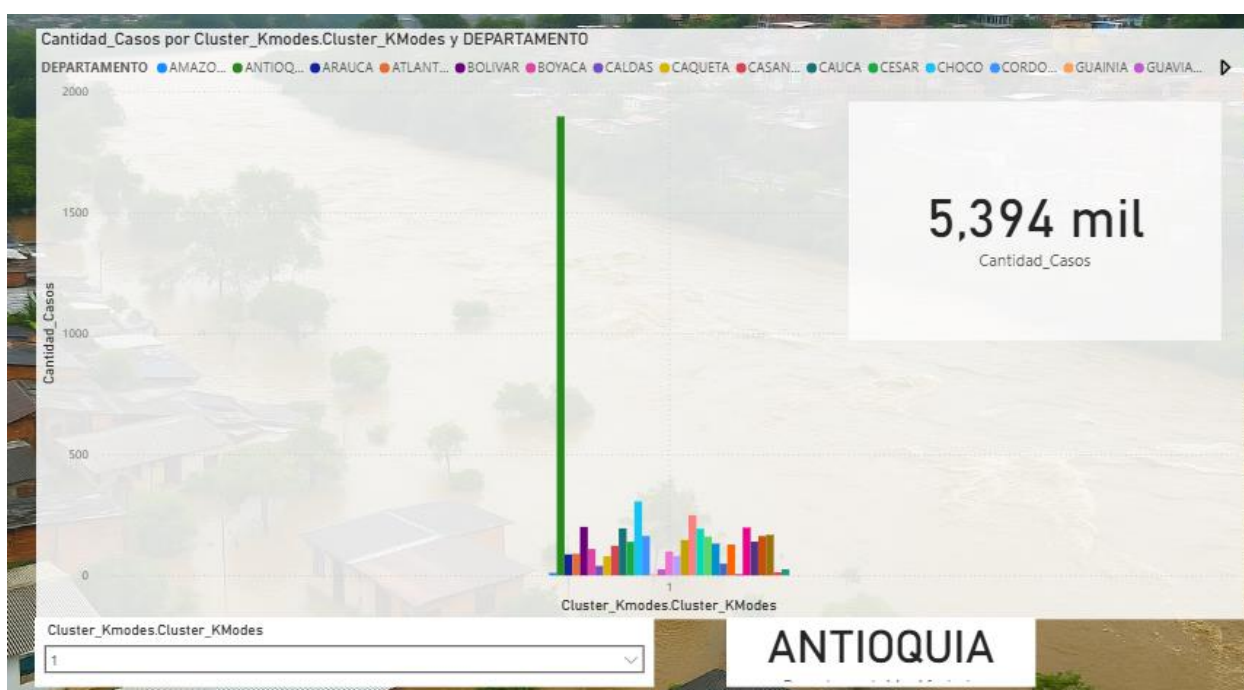


Adicionalmente en la *Figura 15*, se puede realizar un análisis espacial donde se indica que estos eventos del clúster se presentan en *zonas costeras* y *fluviales*, con presencia destacada

en *Antioquia, Atlántico, Chocó, Magdalena, Bolívar y Santander*, así como en otras cuencas hidrográficas del interior. Estas regiones están influenciadas por factores como urbanización acelerada sin planeación, deterioro ambiental y escasa infraestructura de drenaje. Por lo tanto, este clúster apunta a la necesidad de *intervenciones estructurales*, como ampliación de canales, sistemas de alerta temprana y reubicación de asentamientos en zonas de alto riesgo hídrico.

### Figura 15

#### Detalle de Clúster 1



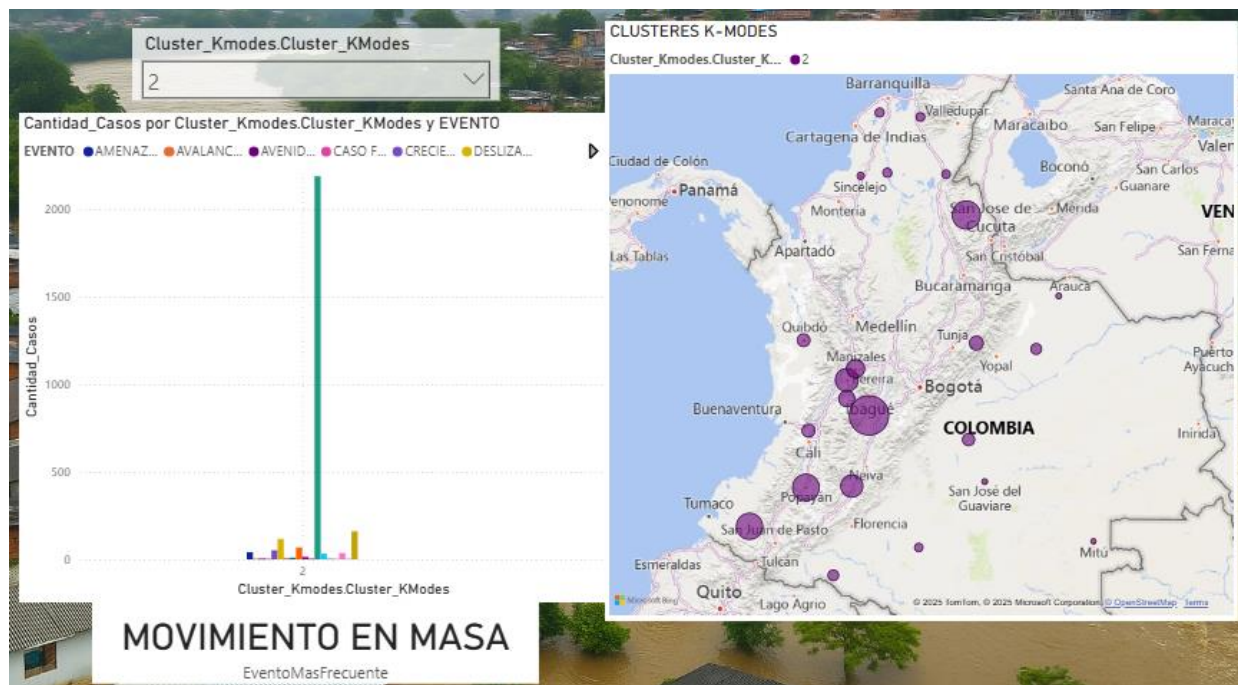
#### Clúster 2 – Movimiento en Masa en el Suroccidente y Oriente

El Clúster 2 agrupa principalmente eventos asociados a la *inestabilidad del terreno*, como *movimiento en masa, aludes y avalanchas*, que suelen desencadenarse después de lluvias persistentes o eventos sísmicos menores. De acuerdo con la *Figura 16*, se puede identificar en la gráfica de barras, que los movimientos en masa son los eventos predominantes, lo que evidencia un patrón físico-geológico propio de las zonas de alta pendiente con suelos erosionables. La alta

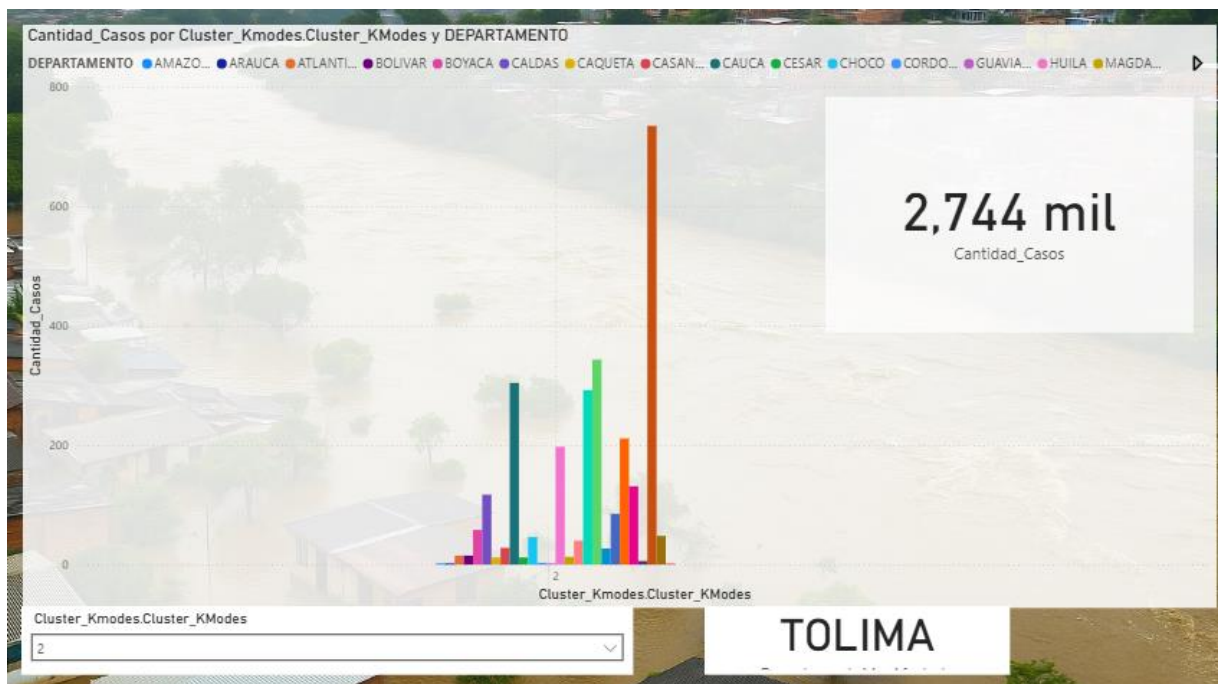
correlación entre este clúster y eventos de remoción en masa sugiere la necesidad de prestar atención a las *condiciones topográficas y de cobertura vegetal* como factores críticos en la generación del riesgo.

**Figura 16**

*Visualización de Clúster 2*



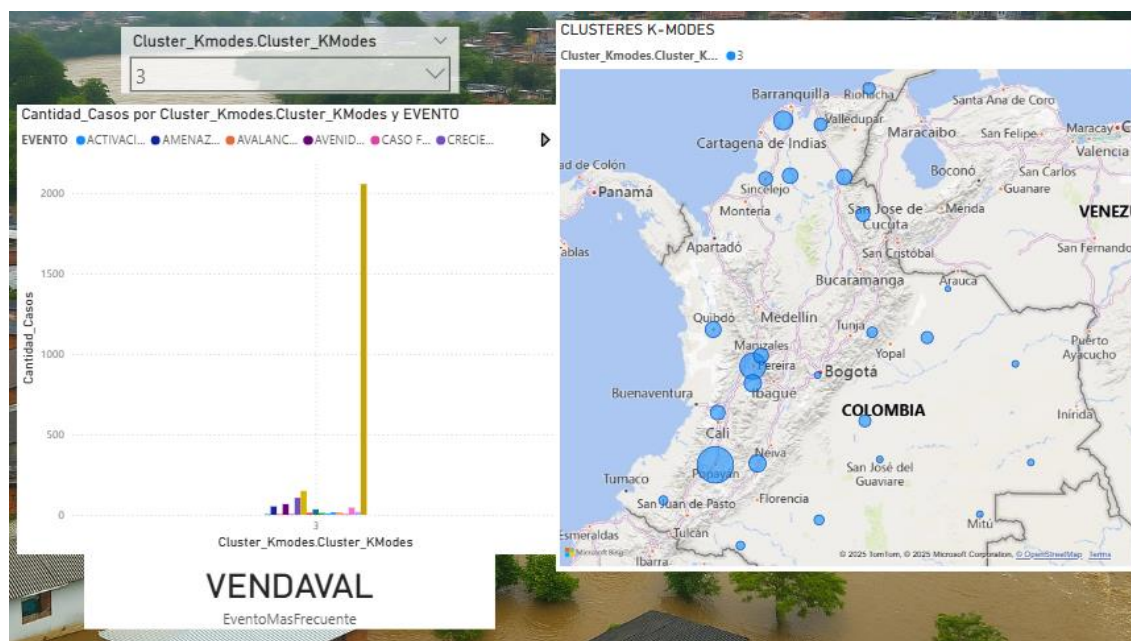
El componente espacial del clúster y acorde a lo que se puede visualizar en la *Figura 17*, esta revela una fuerte presencia en departamentos *andinos como, Tolima, Cauca, Nariño, Huila y Norte de Santander*, regiones caracterizadas por su relieve montañoso, intensas precipitaciones estacionales y procesos de ocupación del suelo en laderas. Este perfil de riesgo requiere estrategias como monitoreo geotécnico, reforestación estratégica, obras de estabilización de taludes y revisión de normativas de uso del suelo. Además, es esencial articular políticas de desarrollo rural sostenible con la prevención del riesgo, dado que muchas de las áreas afectadas están habitadas por comunidades rurales vulnerables.

**Figura 17***Detalle de Clúster 2****Clúster 3 – Vendavales y su Concentración en la Región Caribe y Eje Cafetero***

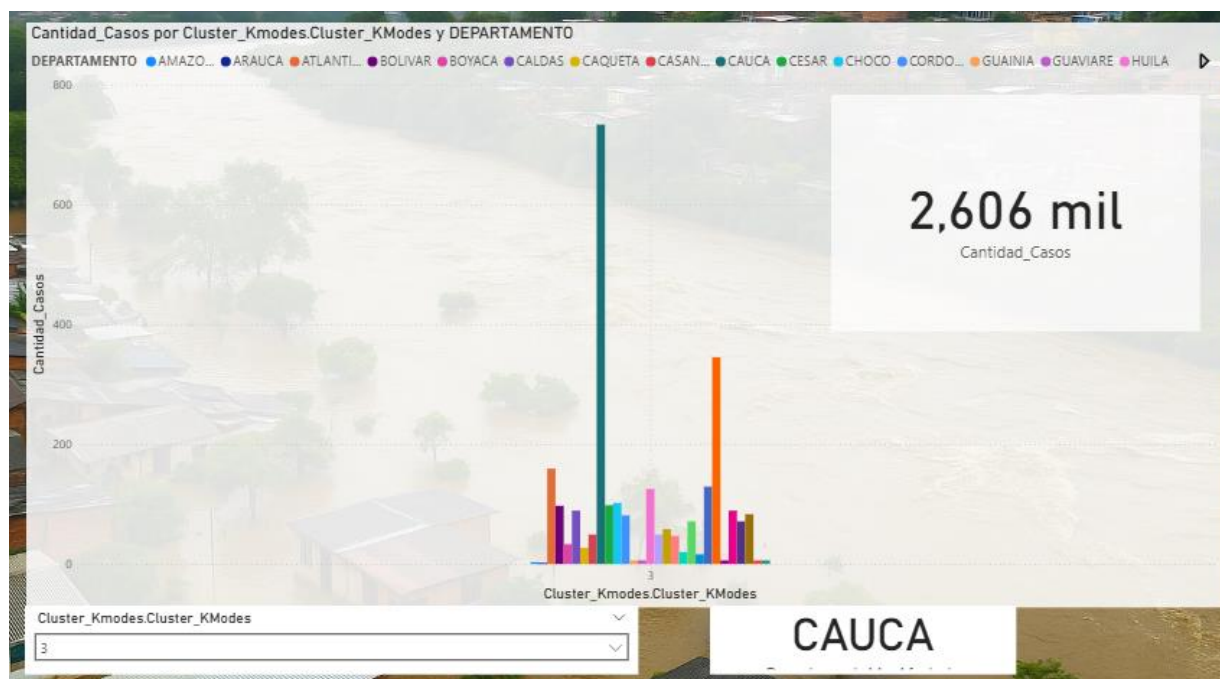
En el Clúster 3, identificado a través del algoritmo no supervisado K-Modes, y apoyado con la *Figura 18*, se evidenció que el evento natural más frecuente es el *vendaval*, el cual presenta una clara dominancia frente a otros tipos de amenazas en esta agrupación. Este comportamiento sugiere que los territorios incluidos en este clúster, localizados principalmente en regiones del centro y suroccidente del país, como el Valle del Cauca, Huila, Tolima y Antioquia, están particularmente expuestos a fenómenos de fuertes vientos. Esta información es clave para la planificación territorial y la gestión del riesgo, ya que permite priorizar estrategias enfocadas en la prevención de daños estructurales por ráfagas de viento, el diseño de sistemas de alerta temprana y la preparación comunitaria frente a este tipo de amenaza específica.

Figura 18

## Visualización de Clúster 3

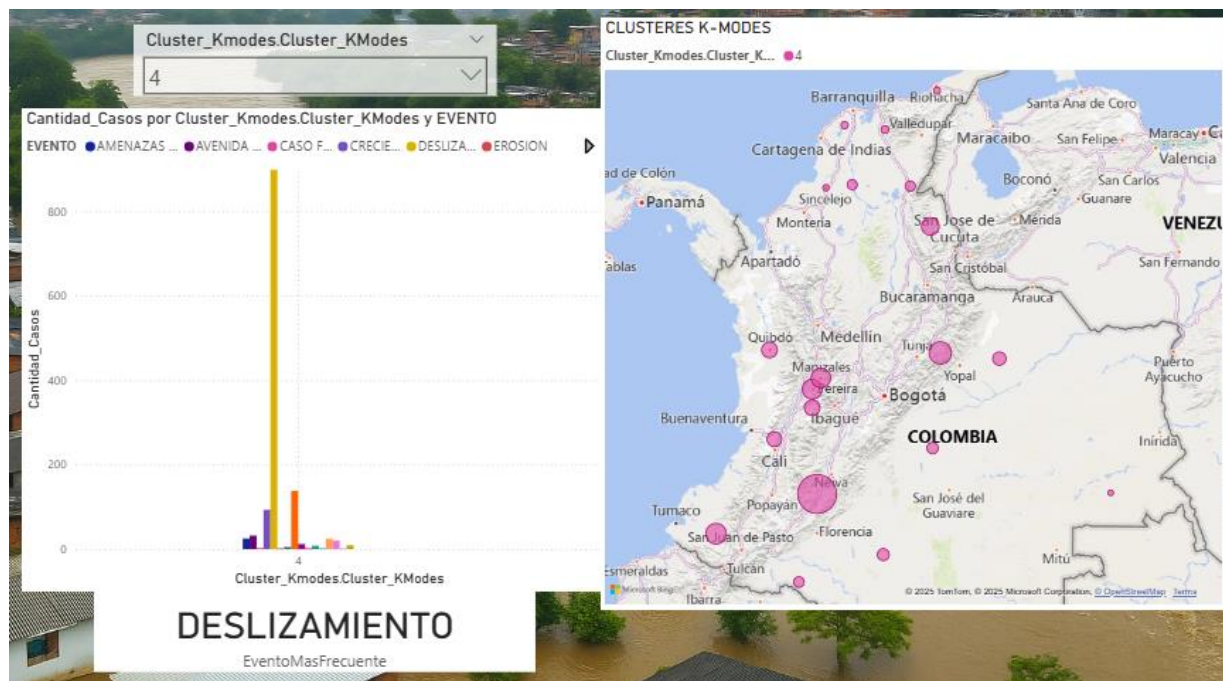


Dentro del Clúster 3 y como se visualiza en la *Figura 19*, el departamento con mayor número de registros de desastres naturales es Cauca, acumulando un total de más de 800 eventos, según lo evidenciado en la visualización. Este resultado resalta la alta exposición del departamento a múltiples amenazas, en especial vendavales, que son el evento predominante en este clúster. La elevada recurrencia de casos en Cauca sugiere condiciones geográficas y climáticas particulares que lo hacen vulnerable, lo cual refuerza la necesidad de priorizar este territorio en los planes de mitigación, atención de emergencias y fortalecimiento de capacidades locales. Este tipo de análisis territorial contribuye a enfocar las acciones de gestión del riesgo de forma diferencial y con base en evidencia.

**Figura 19***Detalle de Clúster 3****Clúster 4 – Deslizamientos y Creciente Súbita en Zonas con Relieve***

De acuerdo con lo que se visualiza en la *Figura 20*, El Clúster 4 está caracterizado por la predominancia del evento natural *deslizamiento*, que se destaca como el tipo de desastre más frecuente dentro de este grupo, tal como se refleja en la visualización de barras. Este fenómeno representa la principal amenaza en las zonas agrupadas en este clúster, evidenciando una recurrencia significativa a lo largo del periodo analizado.

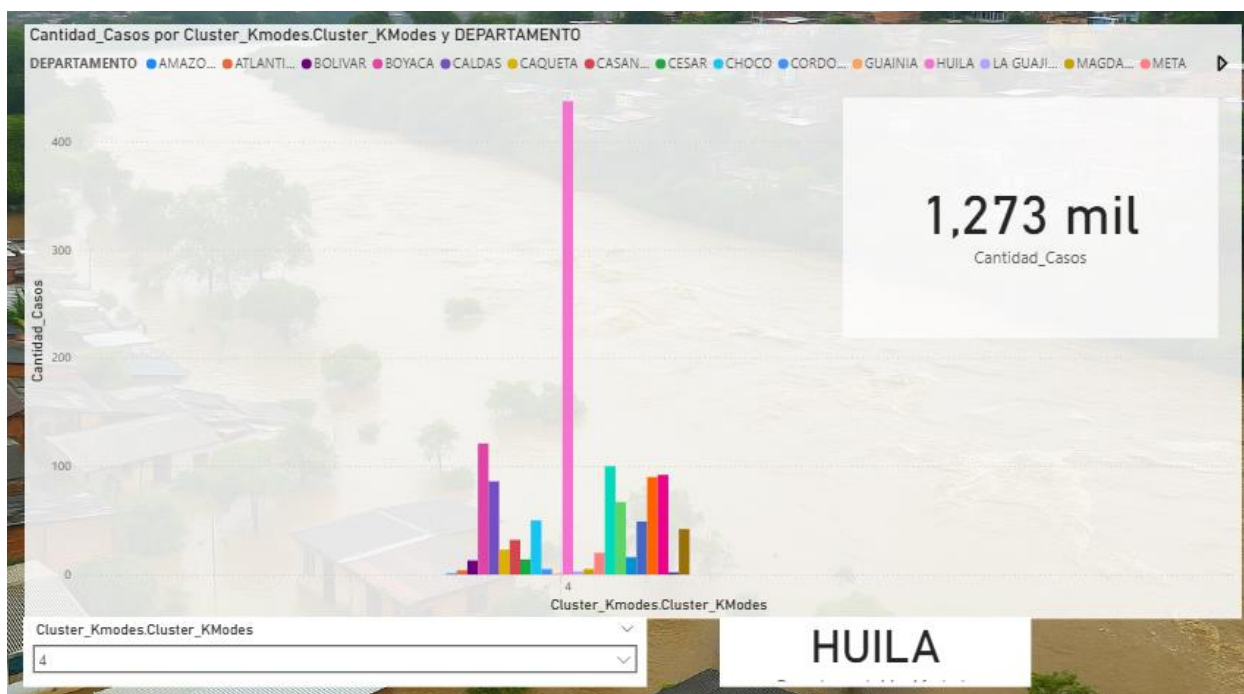
La distribución geográfica, mostrada en el mapa, indica que los municipios y departamentos que conforman este clúster se localizan principalmente en áreas montañosas y de alta pendiente, donde las condiciones geológicas y climáticas favorecen la ocurrencia de deslizamientos. Esto resalta la vulnerabilidad particular de estas regiones ante eventos geodinámicos, que pueden generar impactos significativos en la infraestructura y la población.

**Figura 20***Visualización de Clúster 4*

En el Clúster 4 y de acuerdo como se visualiza en la *Figura 21*, se observa que el departamento con mayor concentración de desastres naturales es Huila, con un total que supera los 400 eventos, lo cual destaca su alta vulnerabilidad dentro de este grupo. Esta cifra refleja la prevalencia significativa de deslizamientos, el evento más frecuente en este clúster, que afecta de manera recurrente a esta región. La concentración de casos en Huila subraya la importancia de enfocar esfuerzos en la gestión del riesgo en esta zona, mediante acciones de prevención, monitoreo geotécnico y fortalecimiento comunitario para reducir el impacto de los movimientos en masa y otros fenómenos asociados.

## Figura 21

### Detalle de Clúster 4



### Comparaciones entre Clústeres

En este apartado se reafirma el enfoque del análisis de clústeres, profundizando en la interpretación de los resultados obtenidos y su importancia para la identificación de áreas especialmente vulnerables a desastres naturales. Se abordarán casos de estudio específicos en los que la aplicación de técnicas de agrupamiento ha contribuido significativamente a la planificación y gestión del riesgo, resaltando cómo estos análisis han respaldado la formulación de políticas públicas y estrategias de mitigación más eficaces.

Asimismo, se realizará una comparación detallada entre los clústeres generados mediante el algoritmo K-Modes, centrándose en dimensiones clave como el tipo de desastre predominante, la ubicación geográfica de los eventos, la frecuencia de ocurrencia, el nivel de impacto potencial y la complejidad asociada a la gestión del riesgo. Esta comparación permitirá identificar patrones

recurrentes, así como contrastes relevantes entre las agrupaciones, facilitando una mejor comprensión del comportamiento regional de los desastres y orientando la toma de decisiones en materia de prevención y respuesta.

### ***Tipo de Desastre Dominante***

De acuerdo con el análisis realizado mediante la técnica de K-Modes, se ha determinado una agrupación de clústeres que permite identificar patrones significativos en la ocurrencia de desastres naturales. Al enfocar la interpretación en el tipo de desastre predominante dentro de cada clúster, se construyó la siguiente tabla resumen, la cual facilita la comprensión de las características principales de cada grupo. Esta clasificación permite no solo entender la naturaleza de los eventos más frecuentes en cada clúster, sino también orientar la formulación de estrategias específicas de mitigación y respuesta ante los diferentes escenarios de riesgo.

**Tabla 4**

*Tipo de Desastre Predominante por cada Clúster*

CLÚSTER	Tipo(S) De Desastre Principal (Es)
0	Incendios forestales
1	Inundaciones + Movimientos en masa
2	Movimientos en masa
3	Vendavales
4	Deslizamientos + Crecientes súbitas

De acuerdo con la *Tabla 4*, que clasifica los tipos de desastre predominantes por clúster, resultado del análisis mediante el algoritmo K-Modes, permite identificar patrones significativos en la ocurrencia de eventos adversos en Colombia. Esta segmentación evidencia que cada clúster

agrupa regiones con características comunes respecto al tipo de amenaza, lo que resulta fundamental para orientar estrategias de mitigación diferenciadas. Por ejemplo, el Clúster 0, asociado principalmente a incendios forestales, sugiere la necesidad de campañas de prevención y control en zonas andinas; mientras que el Clúster 1, que combina inundaciones y movimientos en masa, demanda una gestión integral de riesgos hidrometeorológicos y geotécnicos. Asimismo, los Clústeres 2 y 4, enfocados en movimientos en masa y crecientes súbitas, exigen intervenciones específicas en zonas de alta pendiente y vulnerabilidad geológica. El Clúster 3, por su parte, refleja la concentración de vendavales, lo cual orienta a fortalecer la infraestructura frente a eventos de viento extremo. En conjunto, esta clasificación proporciona una base sólida para focalizar recursos, diseñar políticas más efectivas y fortalecer los sistemas de alerta y planificación territorial con base en las amenazas naturales más recurrentes en cada agrupación.

### ***Ubicación Geográfica Predominante***

La segmentación mediante K-Modes revela patrones geográficos distintivos que permiten una mejor comprensión del comportamiento de los desastres naturales en Colombia. en la *Tabla 5*, se puede observar que el Clúster 0 se caracteriza por una fuerte concentración en la región Andina, una de las zonas más densamente pobladas del país, lo que resalta su importancia estratégica en términos de prevención y mitigación. Por su parte, el Clúster 1 presenta una notable diversidad geográfica al combinar áreas urbanas como Antioquia con regiones selváticas y ribereñas como Chocó y Bolívar, lo que implica desafíos logísticos y ambientales particulares. Finalmente, los Clústeres 2 y 4 comparten una localización predominante en zonas montañosas, las cuales son altamente susceptibles a deslizamientos, lo que subraya la necesidad de estrategias enfocadas en la gestión del riesgo geológico y en infraestructura resiliente. Esta agrupación

permite una aproximación más precisa para la planificación territorial y la respuesta ante emergencias.

**Tabla 5**

*Ubicación Geográfica Predominante por Clúster*

Clúster	Departamentos más Representativos	Región Geográfica Predominante
0	Cundinamarca, Meta, Tolima, Boyacá	Región Andina central
1	Antioquia, Meta, Chocó, Bolívar	Noroccidente y Caribe
2	Norte de Santander, Cauca, Nariño, Huila, Tolima	Montañas del suroccidente y nororiente
3	Cauca, Risaralda, Atlántico, Bolívar	Caribe y Eje Cafetero
4	Huila, Boyacá, Risaralda, Norte de Santander	Zonas montañosas

***Potencial de Impacto y Peligrosidad***

En la *Tabla 6*, se visualiza el potencial de impacto y peligrosidad por clúster, donde se proporciona una visión crítica de los niveles de riesgo asociados a los distintos tipos de desastres naturales identificados en el análisis. Se observa que el Clúster 0, aunque clasificado con un riesgo medio, implica altos costos ambientales y económicos debido a la recurrencia de incendios forestales. El Clúster 1 presenta un riesgo alto, dado que las inundaciones y movimientos en masa afectan directamente a viviendas, infraestructuras viales y vidas humanas, lo cual demanda una atención prioritaria. El Clúster 2 también se considera de alto riesgo, especialmente por la ocurrencia de deslizamientos en zonas rurales, lo que representa una amenaza significativa para comunidades aisladas. En contraste, el Clúster 3, con riesgo medio,

está asociado a vendavales que impactan en mayor medida estructuras vulnerables, mientras que el Clúster 4 destaca como el más crítico, con un riesgo muy alto por su letalidad potencial en áreas densamente pobladas o de difícil acceso. Este análisis permite jerarquizar la urgencia de intervención en cada clúster, fortaleciendo la planificación preventiva y la asignación estratégica de recursos para reducir los impactos adversos en las poblaciones más expuestas.

**Tabla 6**

*Potencial de Impacto y Peligrosidad por Clúster*

Clúster	Riesgo de Víctimas y Daños Materiales	Comentario Relevante
0	Medio	Daños ambientales y económicos altos
1	Alto	Afecta viviendas, vías y vidas humanas
2	Alto	Deslizamientos peligrosos en áreas rurales
3	Medio	Estructuras débiles afectadas por vientos fuertes
4	Muy alto	Riesgo mortal en áreas densamente pobladas o de difícil acceso

***Complejidad para la Gestión del Riesgo***

Acorde a la *Tabla 6*, donde el Clúster 4 representa el mayor desafío en términos de gestión del riesgo, ya que combina una geografía compleja como zonas montañosas o de difícil acceso con una alta letalidad en los eventos registrados. Esta combinación exige intervenciones más sofisticadas, tanto en prevención como en respuesta, incluyendo planes de evacuación robustos y una mejora significativa en la infraestructura crítica. Por otro lado, el Clúster 3 aunque presenta una recurrencia elevada, se caracteriza por ser más predecible y menos severo

en sus consecuencias. Esto permite que pueda ser abordado con relativa facilidad mediante estrategias como la implementación de sistemas de alerta temprana, mantenimiento regular de infraestructuras y campañas de concientización comunitaria. Estas diferencias entre clústeres subrayan la necesidad de adaptar las políticas de gestión de riesgo a las características específicas de cada clúster por lo que se generó la siguiente *Tabla 7*, donde se podrá visualizar el tipo de intervención adecuada para cada tipo de clústeres.

**Tabla 7**

*Complejidad para la Gestión del Riesgo por Clúster*

Clúster	Complejidad en Gestión del Riesgo	Requiere Intervención Multidisciplinaria
0	Media	Prevención comunitaria y monitoreo climático
1	Alta	Coordinación entre entidades urbanas, rurales e hidrológicas
2	Alta	Estudios geotécnicos y obras de mitigación
3	Baja	Infraestructura resistente y alertas tempranas
4	Muy alta	Difícil acceso, intervención rápida limitada

### ***Análisis Comparativo con la NSR-10***

El presente estudio permitió identificar importantes hallazgos sobre la distribución espacial y temporal de los desastres naturales en Colombia, especialmente al analizar los resultados del modelo de clústeres y las visualizaciones dinámicas generadas en Power BI. En este contexto, resulta pertinente realizar una comparación con el marco normativo técnico más relevante del país en materia de construcción resiliente: la *Norma Sismo Resistente Colombiana NSR-10*.

La *NSR-10* establece criterios técnicos para el diseño y construcción de edificaciones en función de la *amenaza sísmica*, clasificando el territorio nacional en zonas de riesgo *bajo*, *intermedio* y *alto* con base en los valores de aceleración pico efectiva. Esta normativa ha sido fundamental para reducir la vulnerabilidad estructural frente a sismos, especialmente en zonas urbanas con alta densidad poblacional. Sin embargo, su alcance es limitado en términos multiamenaza, ya que no contempla de manera integral otros fenómenos naturales altamente recurrentes en Colombia, como los *incendios forestales*, *inundaciones*, *vendavales* o *movimientos en masa*, todos ellos evidenciados como dominantes en el presente análisis.

Al contrastar los hallazgos con el enfoque estructural de la *NSR-10*, se evidencia que existen regiones con baja o intermedia amenaza sísmica según la norma, pero que presentan una alta recurrencia de otros tipos de eventos. Por ejemplo, departamentos como *Cundinamarca*, *Meta*, *Huila* o *Antioquia*, si bien no siempre se ubican en zonas sísmicas críticas, concentran una gran cantidad de emergencias asociadas a otros desastres naturales.

Este contraste pone en evidencia la necesidad de complementar *la NSR-10* con un enfoque territorial multiamenaza, que incorpore datos históricos, climáticos y geológicos para el desarrollo de una normativa más integral y dinámica. Una posibilidad es la formulación de anexos o guías técnicas regionalizadas, que vinculen indicadores de recurrencia de desastres con parámetros de construcción, ordenamiento territorial y planificación del riesgo. Además, la integración de tecnologías como sistemas de información geográfica (SIG) y plataformas de análisis de datos (Power BI, por ejemplo) permitiría adaptar el cumplimiento normativo a las características reales y cambiantes del territorio.

En resumen, la *NSR-10* cumple un papel central en la reducción del riesgo sísmico, pero requiere ser actualizada o complementada con herramientas analíticas y criterios de planificación

basados en datos multiamenaza, como los desarrollados en este proyecto. Este enfoque permitiría avanzar hacia una gestión del riesgo más integral, basada en evidencia y con enfoque preventivo, alineado con los principios de resiliencia territorial y sostenibilidad.

## **Plan para la Optimización de Medidas de Mitigación y Respuesta ante Desastres Naturales**

Con base en el análisis histórico de datos de desastres naturales en Colombia entre los años 2015 y 2022, y en los patrones detectados mediante herramientas analíticas y de visualización en Power BI, se propone un plan integral de optimización de medidas de mitigación y respuesta, enfocado en cinco pilares estratégicos:

### **Fortalecimiento de Sistemas de Monitoreo y Alerta Temprana**

Implementar sensores y estaciones de monitoreo en regiones con alta recurrencia de desastres, especialmente en departamentos como Cundinamarca y Antioquia. La integración de sistemas de información geográfica (SIG), alertas comunitarias en tiempo real y modelos predictivos facilitará una respuesta anticipada frente a eventos como incendios forestales, movimientos en masa e inundaciones.

### **Planeación Territorial Basada en Datos**

Utilizar los resultados del análisis de clústeres para establecer zonas de riesgo prioritarias, promoviendo la formulación de planes de ordenamiento territorial (POT) que consideren la vulnerabilidad ambiental, las amenazas naturales recurrentes y la densidad poblacional. Esto permitirá restringir construcciones en zonas de alto riesgo y diseñar infraestructura resiliente.

### **Educación Comunitaria y Gestión Local del Riesgo**

Desarrollar campañas de concientización ciudadana apoyadas en visualizaciones interactivas, que comuniquen de forma clara los riesgos por región y trimestre. Se deben fortalecer las capacidades locales mediante programas de capacitación en primeros auxilios, evacuación, puntos de encuentro y preparación ante emergencias, con participación activa de juntas de acción comunal y cuerpos de socorro.

### **Optimización de Recursos para Respuesta Rápida**

Diseñar esquemas de asignación dinámica de recursos humanos y materiales basados en la estacionalidad y concentración de eventos. Por ejemplo, reforzar brigadas y equipos de intervención durante los trimestres con mayor presencia de incendios (Q1 y Q3) o de inundaciones (Q2 y Q4), según lo evidenciado en los análisis trimestrales.

### **Evaluación y Retroalimentación Continúa Basada en Visualización**

Utilizar tableros de control Power BI como herramienta de monitoreo constante, para hacer seguimiento a la ocurrencia de eventos, efectividad de las respuestas y evolución de los riesgos. Esto permitirá ajustar estrategias de forma oportuna y con evidencia empírica, fomentando un ciclo de mejora continua en la gestión del riesgo.

## Conclusiones

El presente estudio demostró que el uso de herramientas de análisis y visualización de datos, como Power BI, permite no solo representar gráficamente la distribución histórica de desastres naturales en Colombia, sino también identificar patrones espaciales, temporales y categóricos que son fundamentales para la toma de decisiones estratégicas en gestión del riesgo.

Entre los hallazgos más relevantes se destaca la identificación de los departamentos de *Cundinamarca* y *Antioquia* como zonas de alta incidencia, siendo los *incendios forestales* y los *movimientos en masa* los eventos más recurrentes y peligrosos. Asimismo, el análisis por trimestres evidenció picos estacionales significativos, particularmente en los trimestres primero y tercero para incendios, y segundo y cuarto para inundaciones y deslizamientos.

La segmentación por clústeres mediante el algoritmo K-Modes permitió categorizar el territorio nacional en zonas homogéneas de riesgo, revelando agrupamientos de amenazas específicas y su distribución geográfica predominante. Este enfoque no solo aporta valor desde el punto de vista analítico, sino que se convierte en una herramienta clave para orientar políticas públicas diferenciadas.

Por otra parte, resulta pertinente hacer referencia a la *Norma Sismo Resistente Colombiana (NSR-10)*. Aunque esta norma establece lineamientos técnicos para el diseño y construcción de edificaciones seguras frente a amenazas sísmicas, su aplicación sigue siendo limitada en ciertas zonas rurales o de expansión urbana informal. A partir de los hallazgos de este estudio, se sugiere que *la NSR-10* podría fortalecerse incluyendo lineamientos regionalizados basados en la ocurrencia histórica de múltiples amenazas, no solo sísmicas, como deslizamientos, inundaciones o incendios forestales. Asimismo, se hace necesario integrar los

datos territoriales en los procesos de licenciamiento y planificación urbana, especialmente en zonas identificadas como de alta concentración de desastres.

Finalmente, la implementación de un sistema de visualización interactiva no solo mejora la comprensión de los datos, sino que facilita la comunicación efectiva con tomadores de decisiones, organismos de emergencia y comunidades. Este proyecto demuestra el potencial de la analítica de datos como instrumento fundamental para *optimizar medidas de mitigación, fortalecer capacidades de respuesta y construir territorios más resilientes ante desastres naturales.*

## Referencias

- Abdollahi, A. &. (2023). *Explainable artificial intelligence (XAI) for interpreting the*. Science of The Total Environment.
- Ahmad, K. K. (2023). FireXnet: An explainable AI-based tailored Deep learning model for wildfire detection on resource-constrained devices.
- Arjona, M. A. (2017). Estudio comparativo de técnicas de minería de datos para la predicción de rutas de huracanes. *Tecnología Educativa Revista CONAIC*.
- Ayala, J. &. (2023). Desastres naturales en Colombia: un análisis regional.
- Calderón Ramírez, D. &. (2017). *El ordenamiento territorial para la gestión del riesgo de desastres en Colombia*. Obtenido de <http://dx.doi.org/10.12804/revistas.urosario.edu.co/territorios/a.4795>
- Chen, H. C. (2012). *Business intelligence and analytics: From big data to big impact*. *MIS Quarterly*,. Obtenido de <https://doi.org/10.2307/41703503>
- Cutter, S. L. (2008). *Disaster resilience indicators for benchmarking baseline conditions*. *Journal of Homeland Security and Emergency Management*.
- Díaz, R. &. (2019). Análisis predictivo de activos mineros para obtención de intervalo de falla mediante algoritmos de machine learning.
- Gutiérrez, R. Y.-L.-A.-W. (2023). *Vulnerabilidad ante desastres naturales: una visión desde el enfoque socio-comunicacional Algunas consideraciones sobre gestión de riesgos naturales en Perú*. Maracaibo, Venezuela: Universidad Nacional Experimental Rafael María Baralt.
- Han, J. K. (2011). *Data Mining: Concepts and Techniques*. Morgan Kaufmann.
- IDIGER. (2024). *idiger.gov.co*. Obtenido de <https://www.idiger.gov.co/rsismico>

- Levy, J. K. (2005). *Promoting disaster-resilient communities: The great Sumatra–Andaman earthquake of 26 December 2004 and the resulting Indian Ocean tsunami*. Disaster Prevention and Management: An International Journal.
- Macías, Y. Z. (2024). Revisión sistemática de la literatura sobre técnicas de inteligencia artificial en la predicción de desastres naturales. *RevistaG-ner@ndo V°5*, págs. 1205–1235.
- Microsoft. (2023). <https://learn.microsoft.com/en-us/power-bi/>. Obtenido de <https://learn.microsoft.com/en-us/power-bi/>
- Talha Ahmed Khan, M. M. (2020). Reconocimiento previo de inundaciones repentinas: análisis concreto de la configuración óptima de la red neuronal para la detección de múltiples resoluciones. IEEE Access.
- Yaseen, M. W. (2023). Artificial Intelligence Based Flood Forecasting for River Hunza at Danyor Station in Pakistan. Sciendo.
- Unidad Nacional para la Gestión del Riesgo de Desastres. (2023). *Informe de gestión UNGRD 2023*. <https://portal.gestiondelriesgo.gov.co/Documents/Informes-de-Gestion/INFORME-DE-GESTION-UNGRD-2023.pdf>

## Apéndices

### Apéndice A

Enlace Plataforma Power BI



Nota. se puede visualizar en <https://app.powerbi.com/view?r=eyJrIjoj%20ZDVkMDVjODItYThiYy00ZTY5LTkyYmItNWViOTJiZTBINzdmIiwidCI6ImZjMDA1NDdhLTl0YmItNGU0Zi05ZDYxLTczZmNhNWViOWRmMyIsImMiOjR9>

0ZDVkMDVjODItYThiYy00ZTY5LTkyYmItNWViOTJiZTBINzdmIiwidCI6ImZjMDA1NDdh

LTl0YmItNGU0Zi05ZDYxLTczZmNhNWViOWRmMyIsImMiOjR9