



ANÁLISIS DE DATOS

DESASTRES NATURALES

IMPACTO HUMANO Y ECONÓMICO

Soluciones aplicadas de Análisis de Datos y Aprendizaje Automático para entender y prever la magnitud destructiva de fenómenos naturales, con el fin de mitigar su impacto.

Equipo 2
Alejandra Eng
Ángel Bardon
Elena Blanco
Rafael De Marco



OBJETIVOS

Académicos:
Adquirir conocimientos y habilidades de análisis de datos en un contexto real, mediante metodología CRISP-DM en el tema sobre desastres naturales.

Aplicación Práctica:

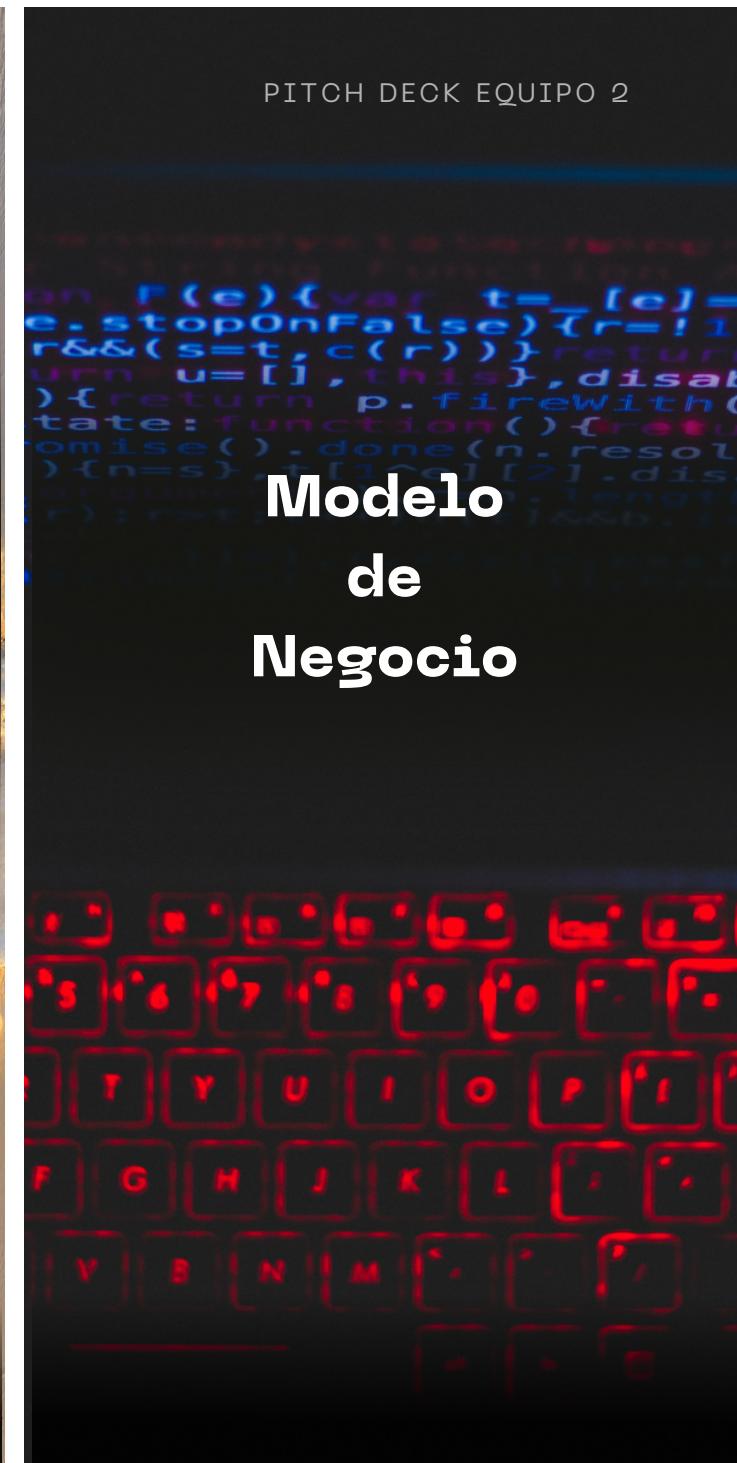
Conocer la magnitud y el impacto socio-económicos para evaluar riesgos y diseñar soluciones prácticas como consultoría de inteligencia de negocio, diseño de dashboards y modelos predictivos.



EDA:
Terremotos
Sequías
Tormentas
Inundaciones



Modelo Predictivo:
Nivel de Destrucción



Modelo de Negocio

DESASTRES NATURALES

Base de datos

EM-DAT (Emergency Events Database)

Fuente: Centre for Research on the Epidemiology of Disasters (CRED), Univ. Católica de Lovaina (Bélgica).

Ocurrencia de desastres masivos desde 1900 hasta la actualidad.

<https://www.emdat.be/>

<https://www.emdat.be/guidelines>

Fuente



Excel

Número de registros

26.004 filas

50 columnas

Grupos de desastres

Naturales

Tecnológicos
Complejos

Selección

16.636 filas

26 columnas

Período que abarca

122 años

1900 a 2023

Tipos de desastres

15 4

Inundación

Sequía

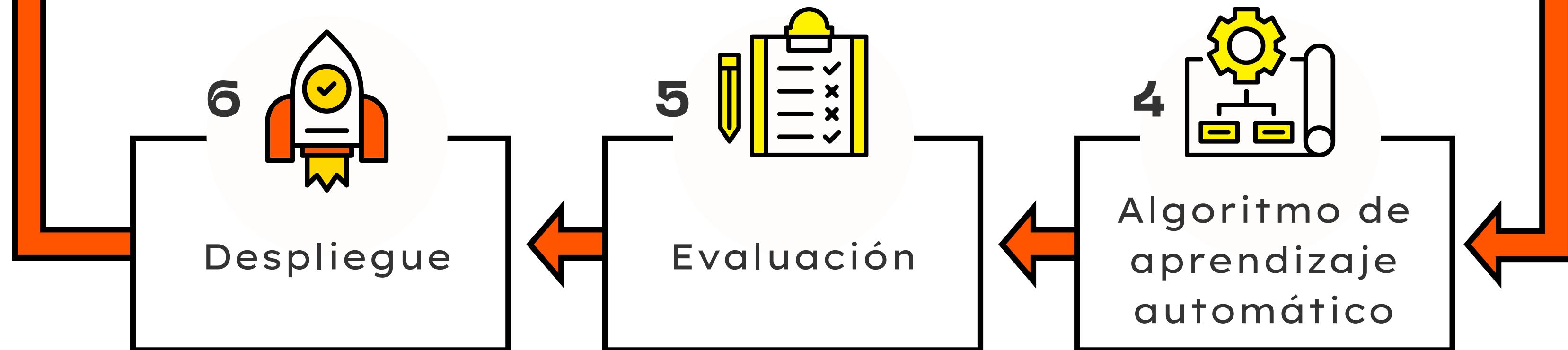
Terremoto

Tormenta

DESASTRES NATURALES

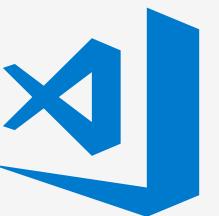


METODOLOGÍA CRISP-DM

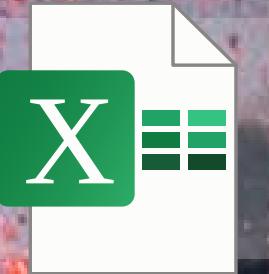


DESASTRES NATURALES

PROCESO Y HERRAMIENTAS



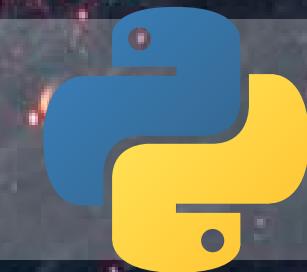
Comprensión del problema



Entendimiento de los datos



Preparación de datos



Modelado

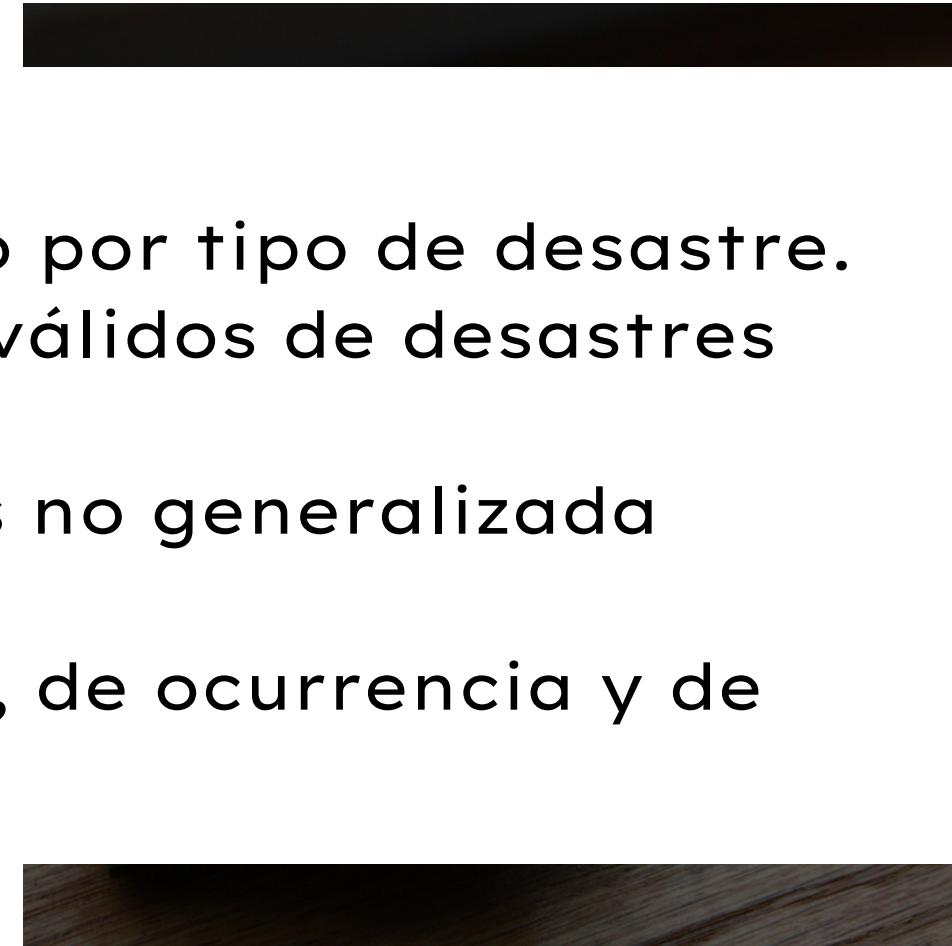


Business Intelligence



PROCESOS GENERALES EDA

- Selección de 4 tipos de desastres.
- Limpieza y segmentación con filtro por tipo de desastre.
- Outliers no se eliminan: son datos válidos de desastres con grandes impactos.
- Limpieza y relleno de valores nulos no generalizada debido a naturaleza del dataset.
- Análisis de categorías geográficas, de ocurrencia y de impacto.



BUSINESS INTELLIGENCE

Gestión del socorro en casos de desastre



DISASTER RELIEF
BI Dashboard



ANÁLISIS DE DATOS. EQUIPO 2

Herramientas





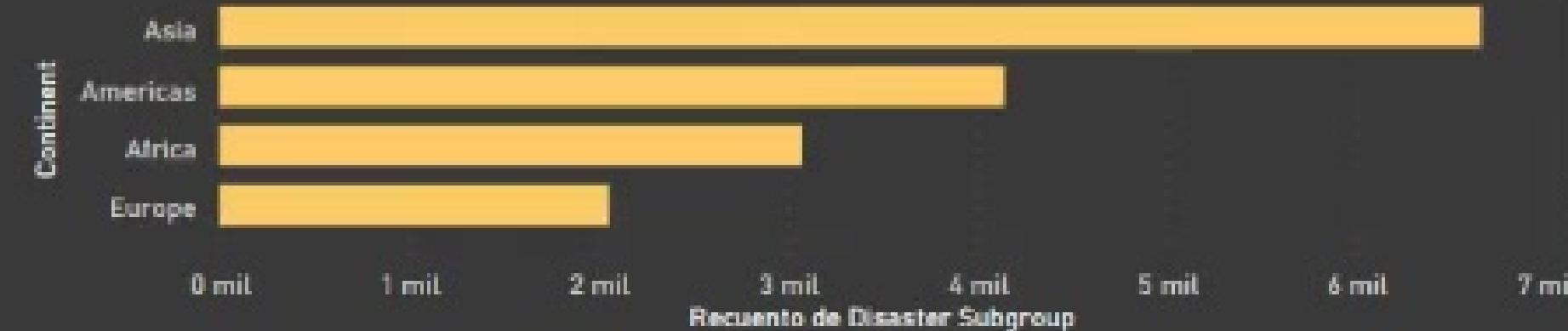
Desastres Naturales

Subgrupos y Datos Generales

Fuente: EM-DAT con datos desde 1900-2023.

Clasificación original de desastres naturales:
Meteorológica, Hidrológica, Geofísica, Climatológica y Biológica.

Por Continente



Tipos de Desastres dentro de Subgrupos de Desastres

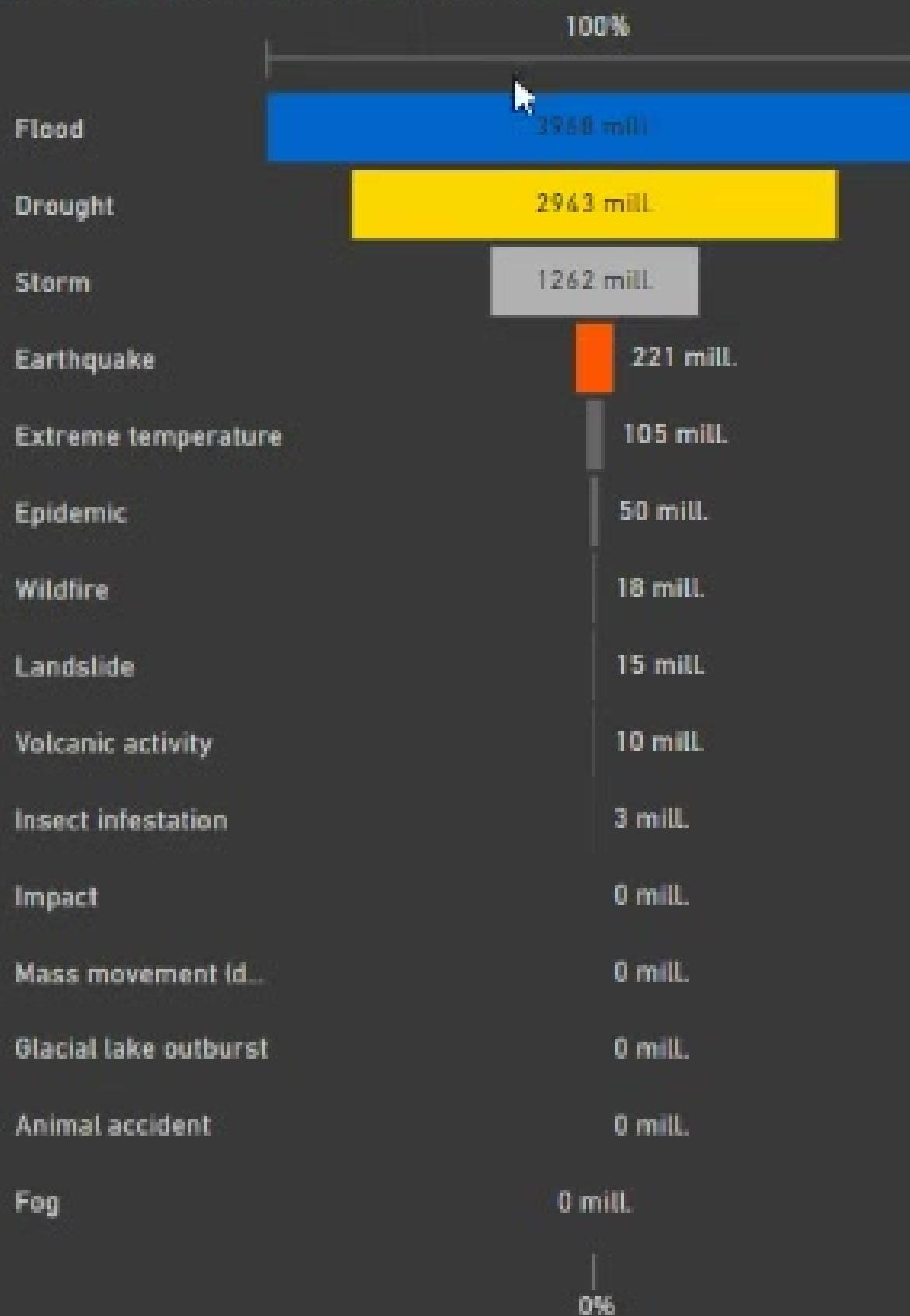




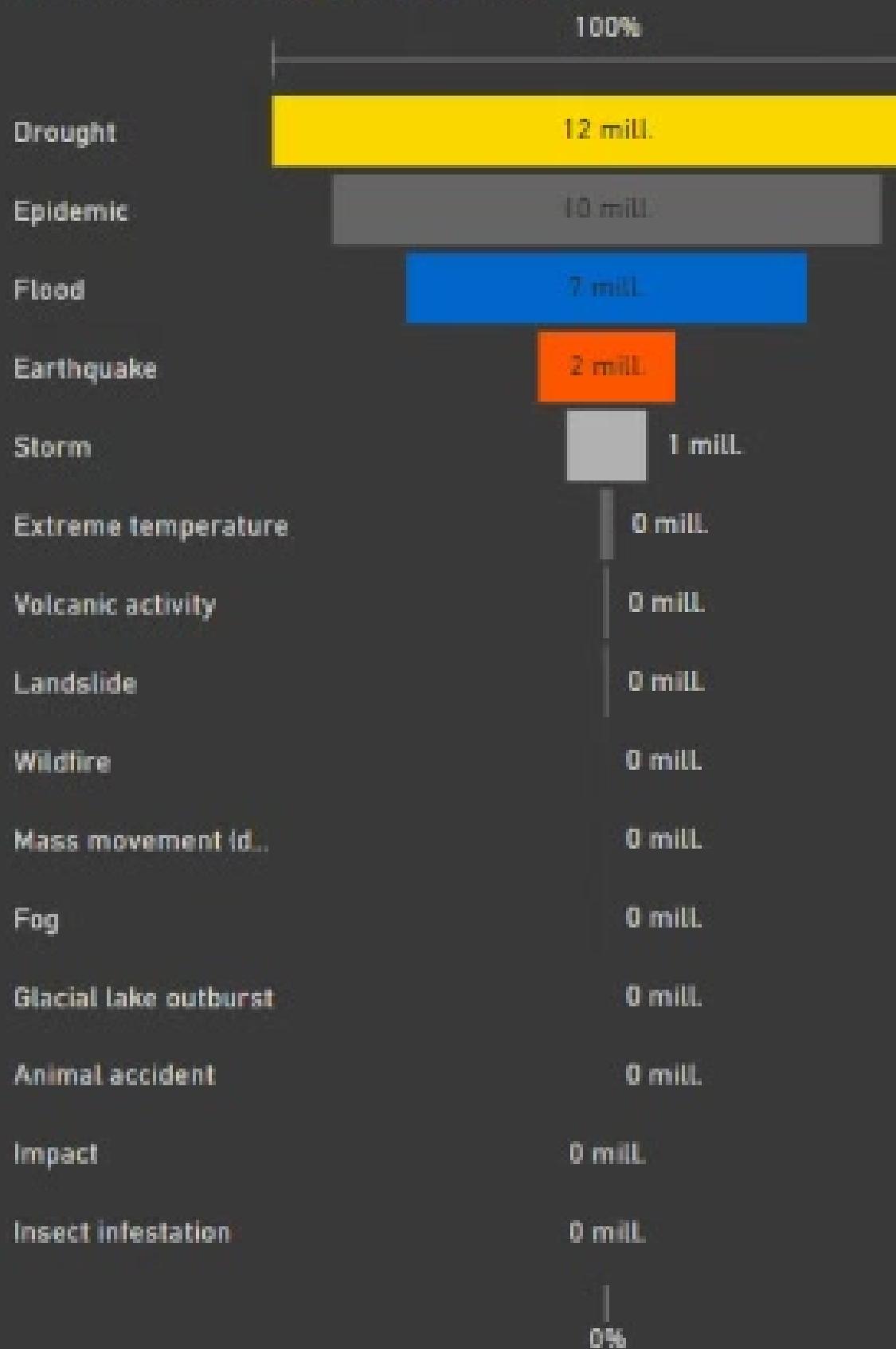
Desastres Naturales

Impactos por Tipo

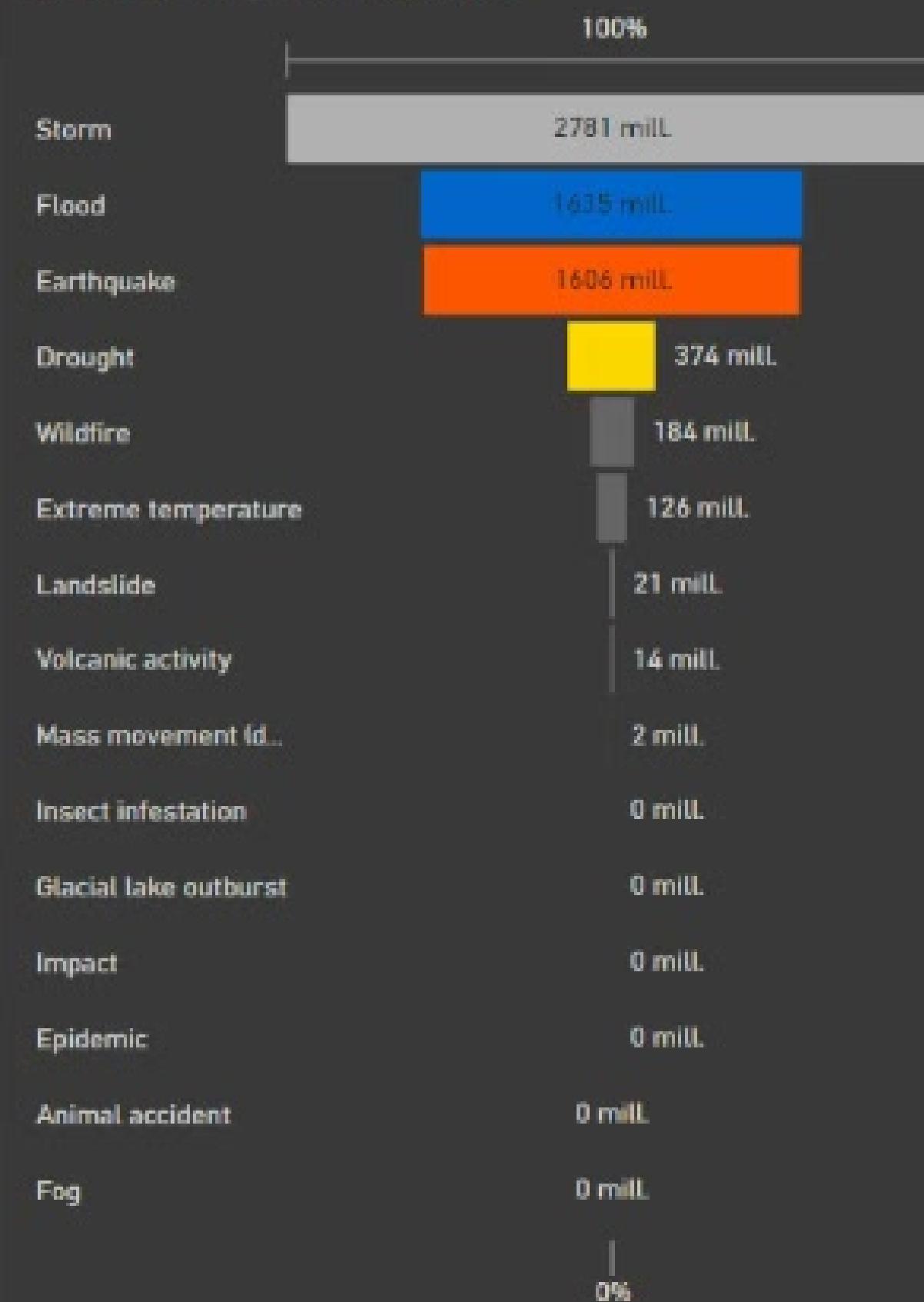
Afectados por Tipo de Desastre



Fallecidos por Tipo de Desastre



Daños por Tipo de Desastre

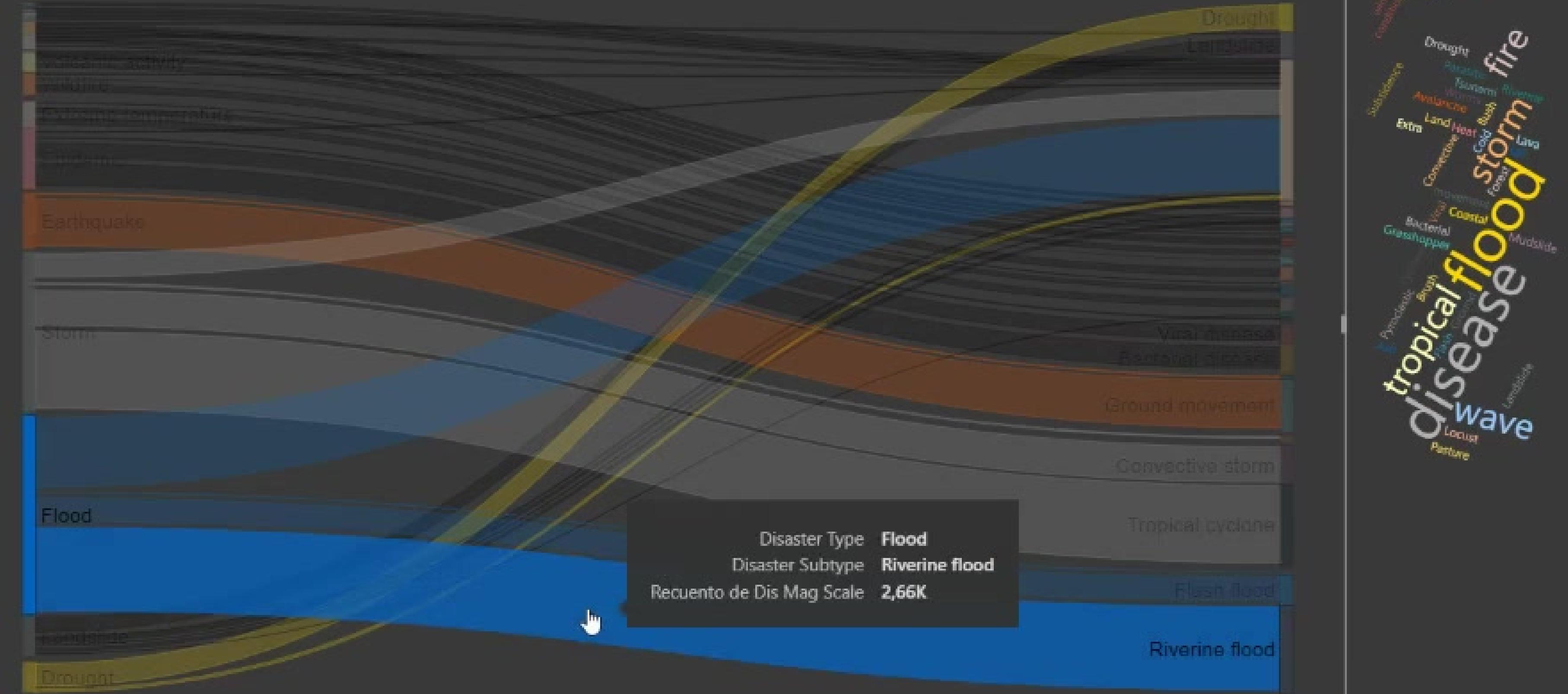




Desastres Naturales

Tipos y Subtipos

Diagrama Sankey: representación del flujo de tipos a subtipos.



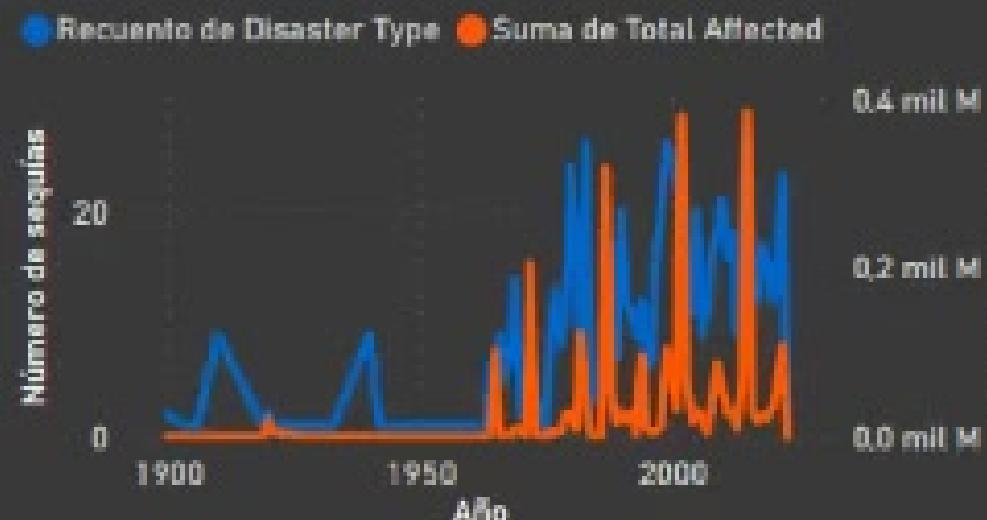


Desastres Naturales

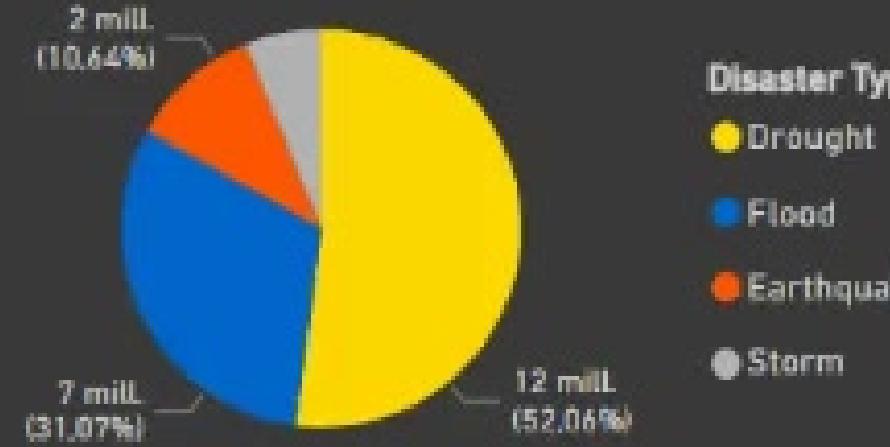
Sequías

10

Análisis temporal por frecuencia y total de afectados



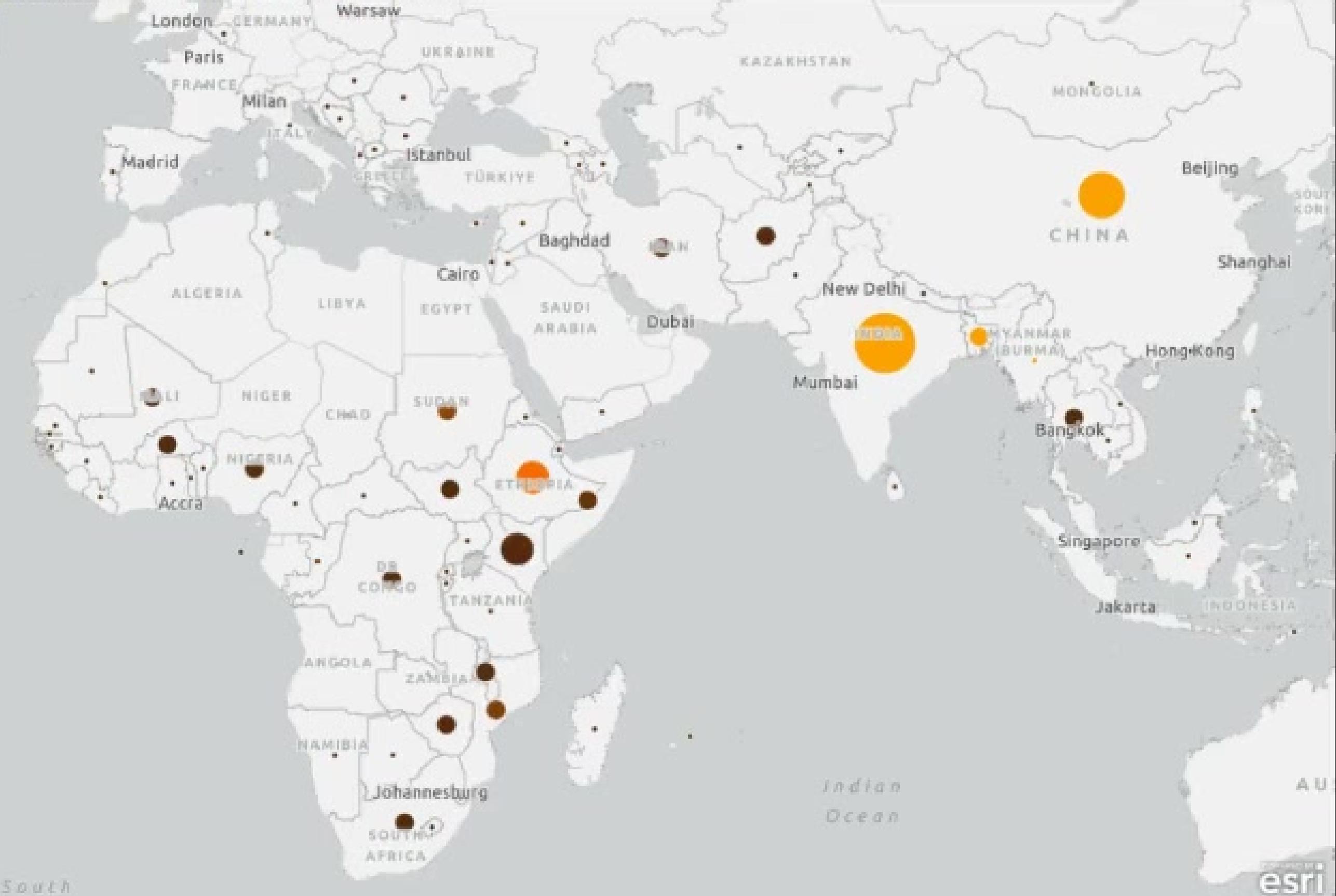
Fallecidos por tipo de desastre



Frecuencia de sequías por continente



Total de Afectados y Total de Fallecidos por País



Frecuencia de sequías por países



Disaster Type • Drought

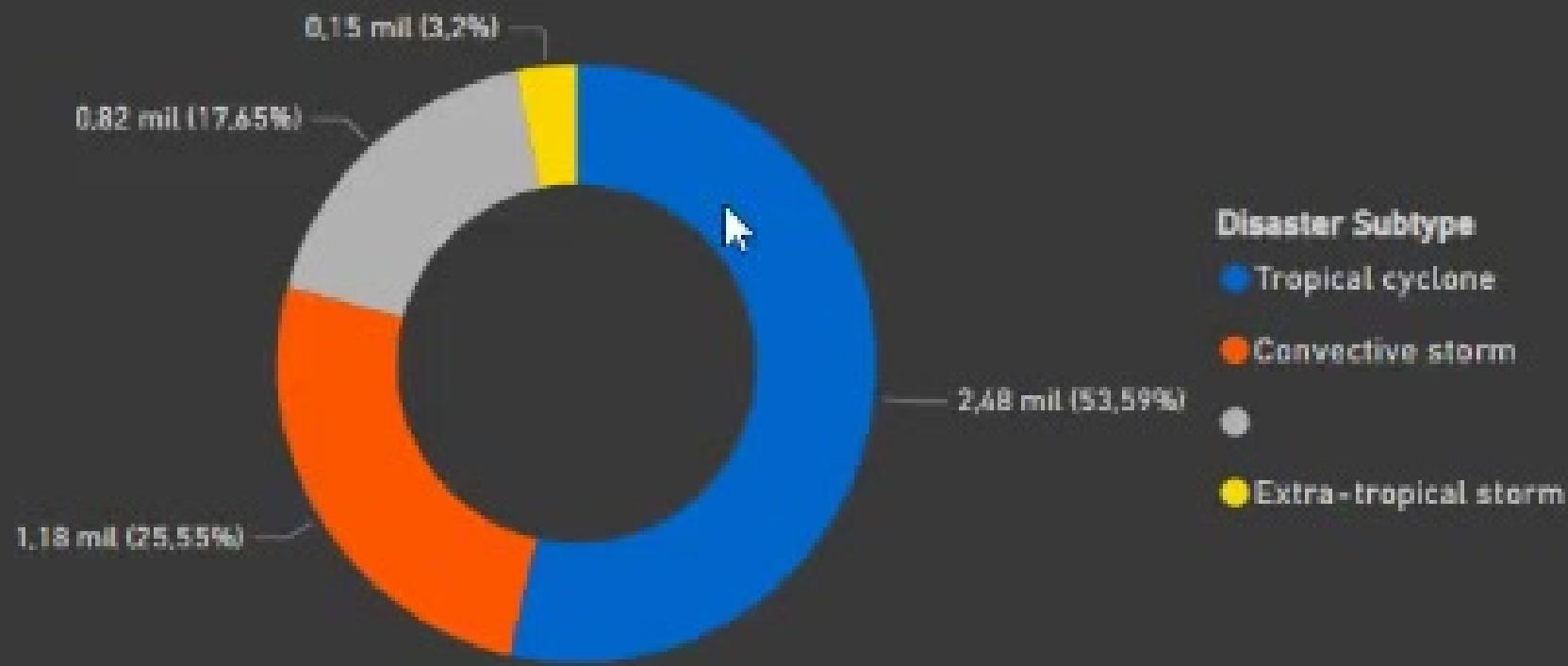




Desastres Naturales

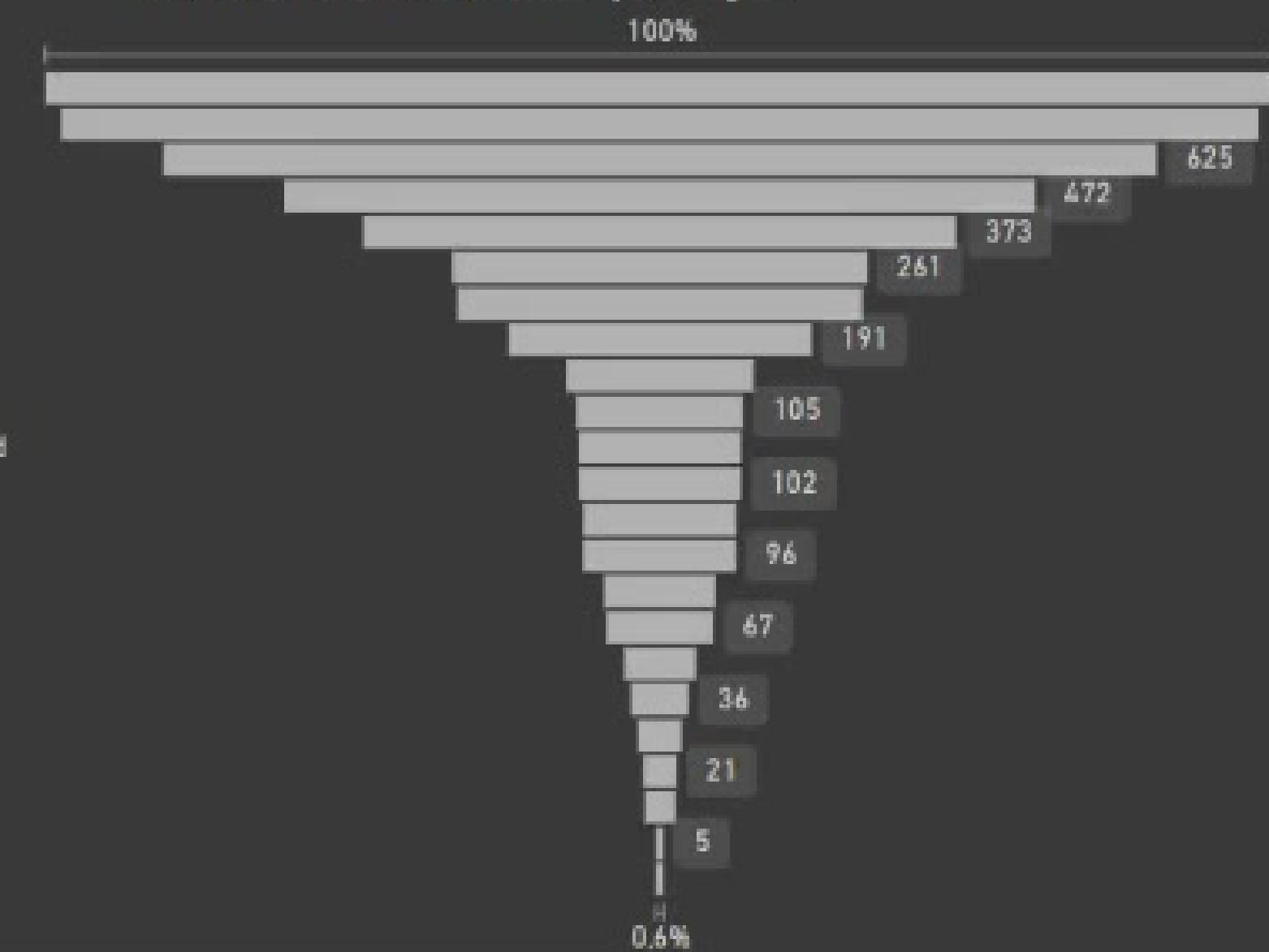
Tormentas

Distribución por subtipos de tormenta

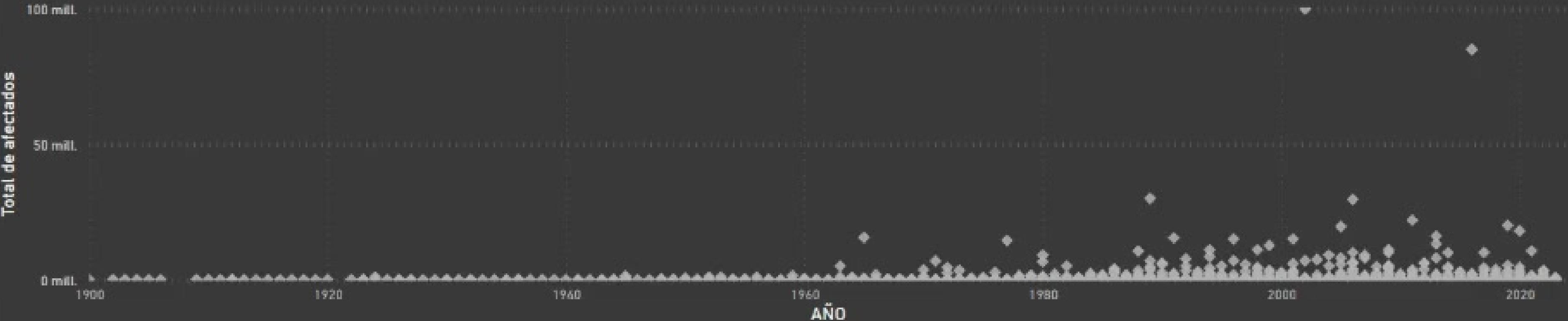


Northern America
Eastern Asia
South-Eastern Asia
Southern Asia
Caribbean
Western Europe
Central America
Eastern Africa
Melanesia
Northern Europe
Australia and New Zealand
Eastern Europe
South America
Southern Europe
Polynesia
Western Asia
Southern Africa
Western Africa
Micronesia
Middle Africa
Northern Africa
Central Asia
Russian Federation

Frecuencia de tormentas por región



Evolución temporal del número de afectados





Desastres Naturales

Terremotos



Magnitud por Año

Dis Mag Value ● 1 ● 3 ● 4 ● 5 ● 6 ● 7 ● 8 ● 9 ● 10

20

15

10

5

0

Recuento de Dis Mag Value

1900

1920

1940

1960

1980

2000

2020

Year

Terremotos a través de los años

Count

1

1980

1990

2000

2010

2020

Magnitud del Evento



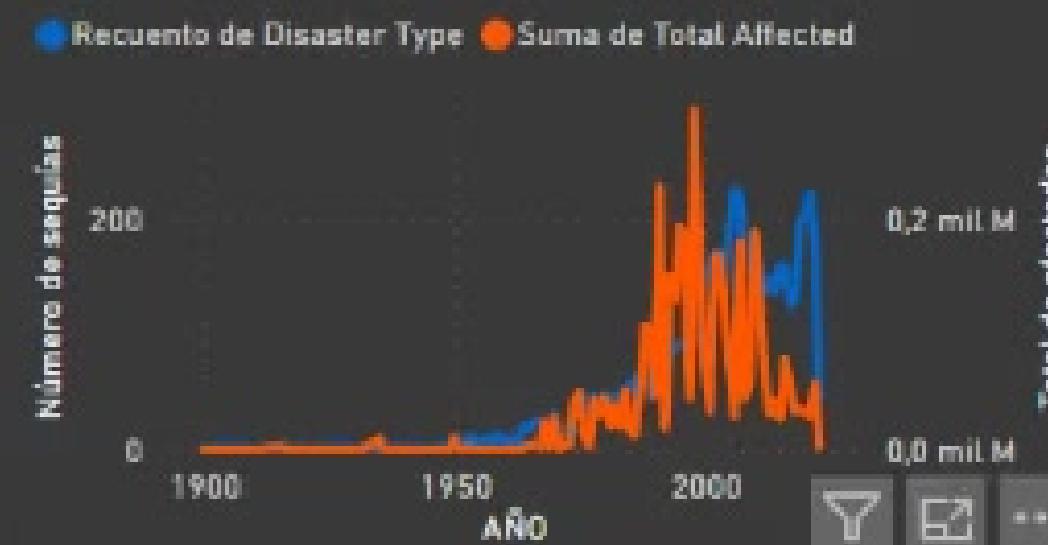
Desastres Naturales

Mapa de Geolocalización de Terremotos

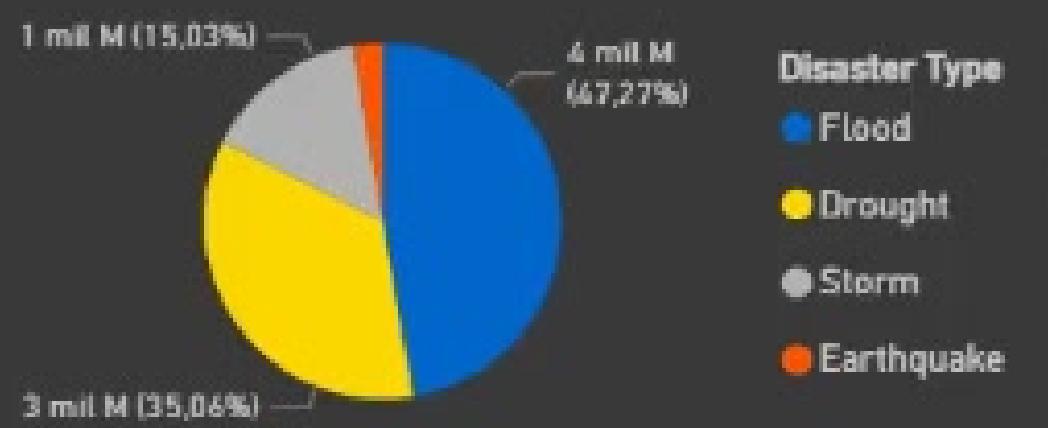




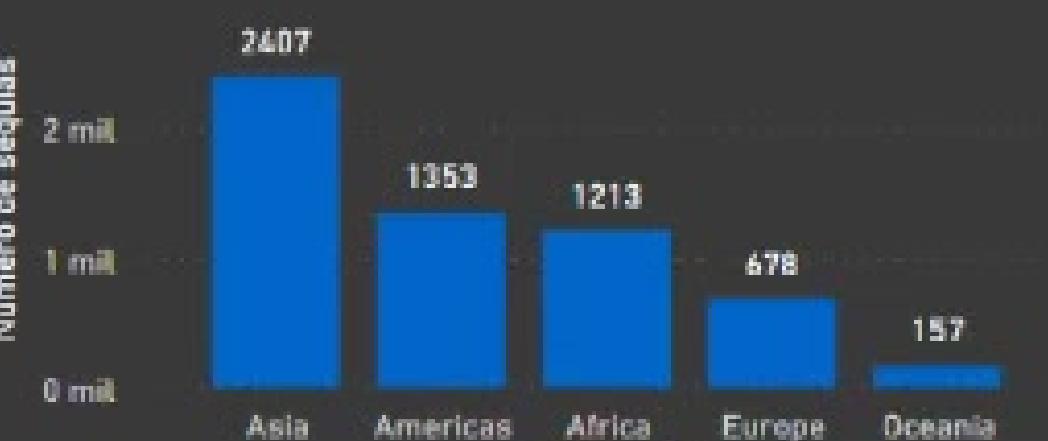
Análisis temporal por frecuencia y total de afectados



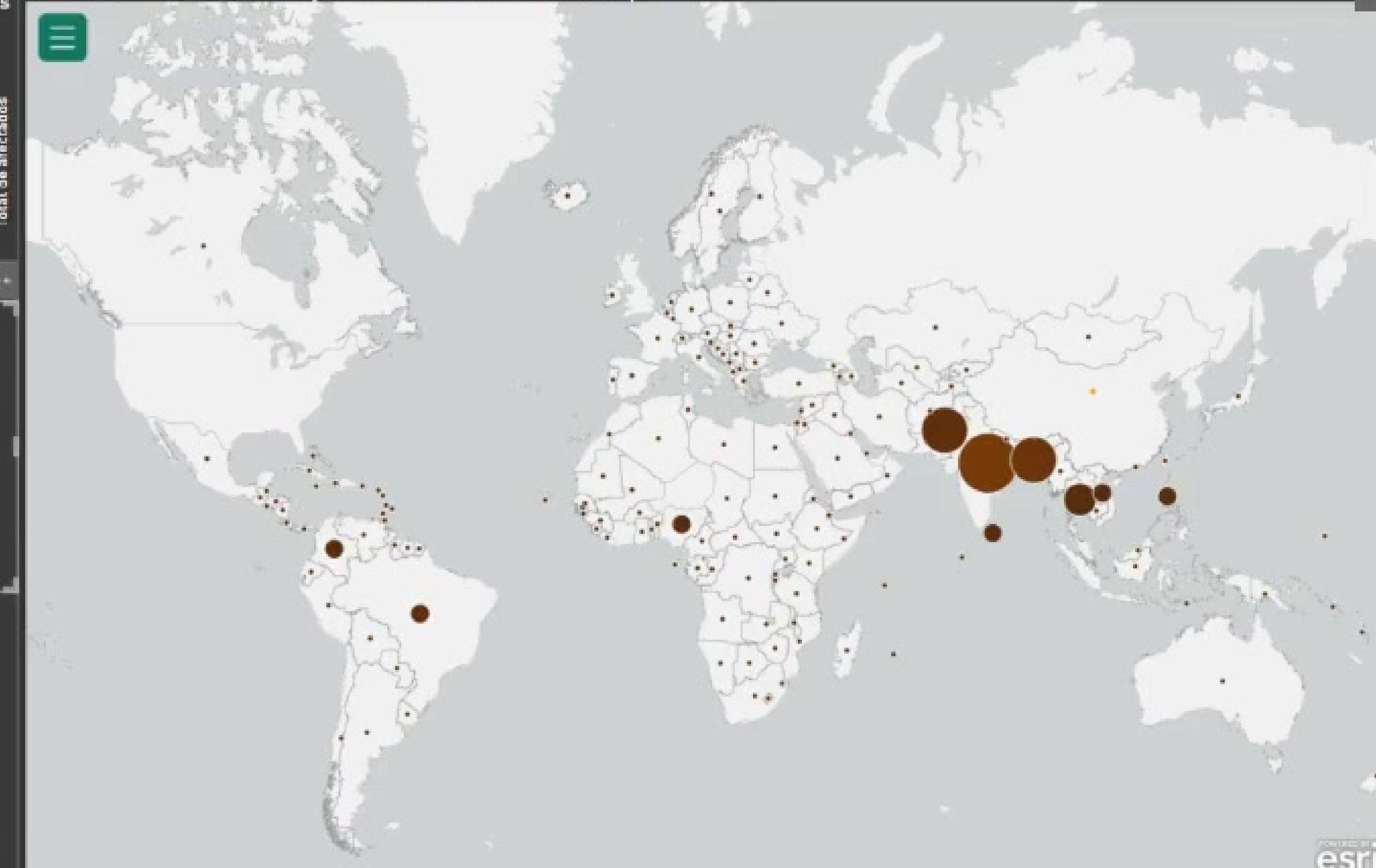
Fallecidos por tipo de desastre



Frecuencia de inundaciones por continente



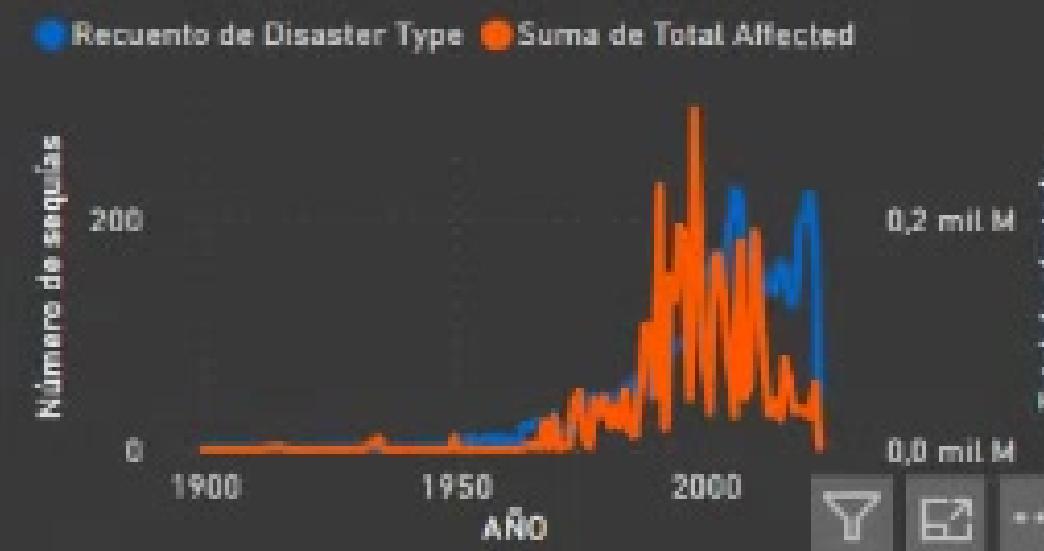
Suma de Total Afectados y Suma de Total de Muertes por País



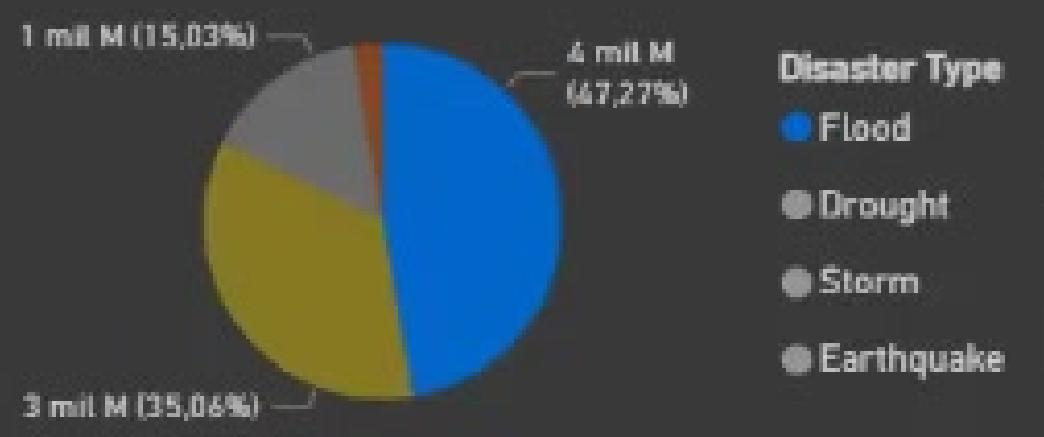
Desastres Naturales

Inundaciones

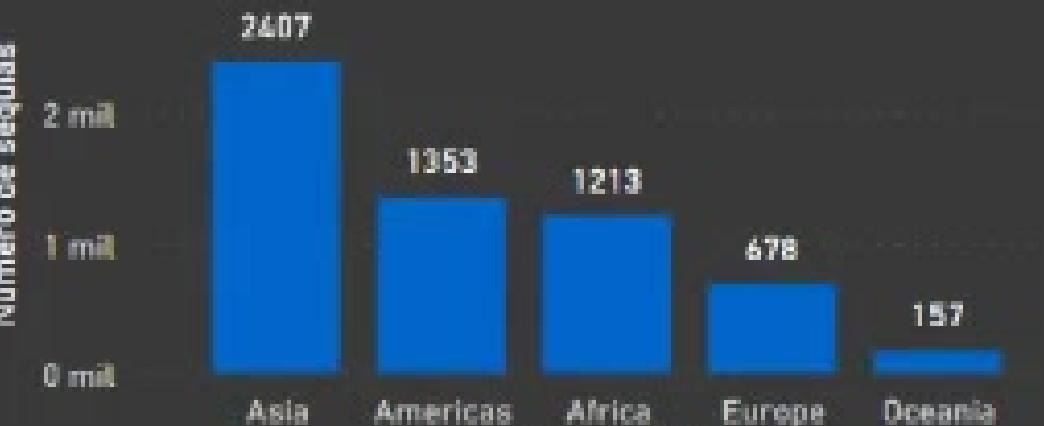
Análisis temporal por frecuencia y total de afectados



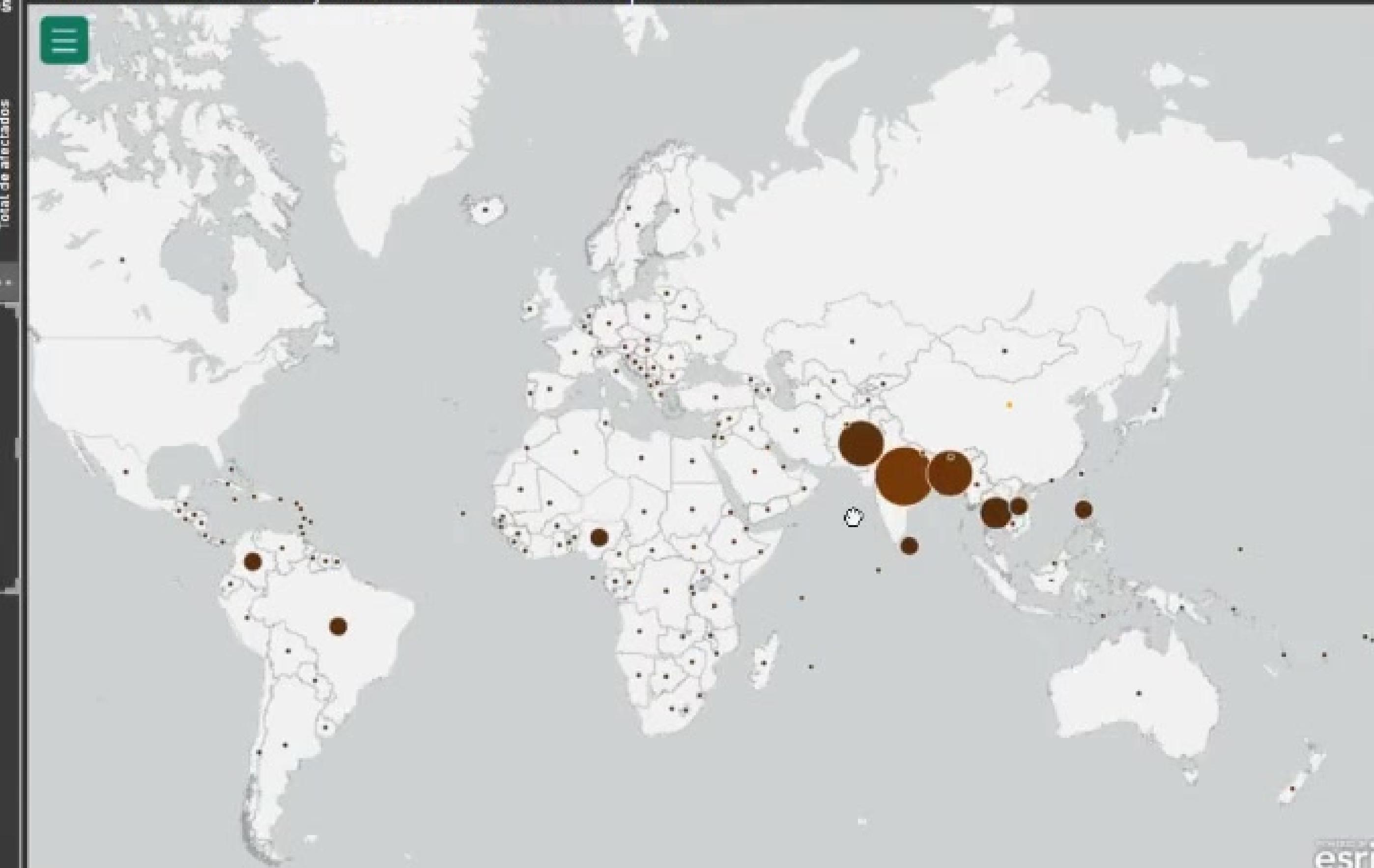
Fallecidos por tipo de desastre



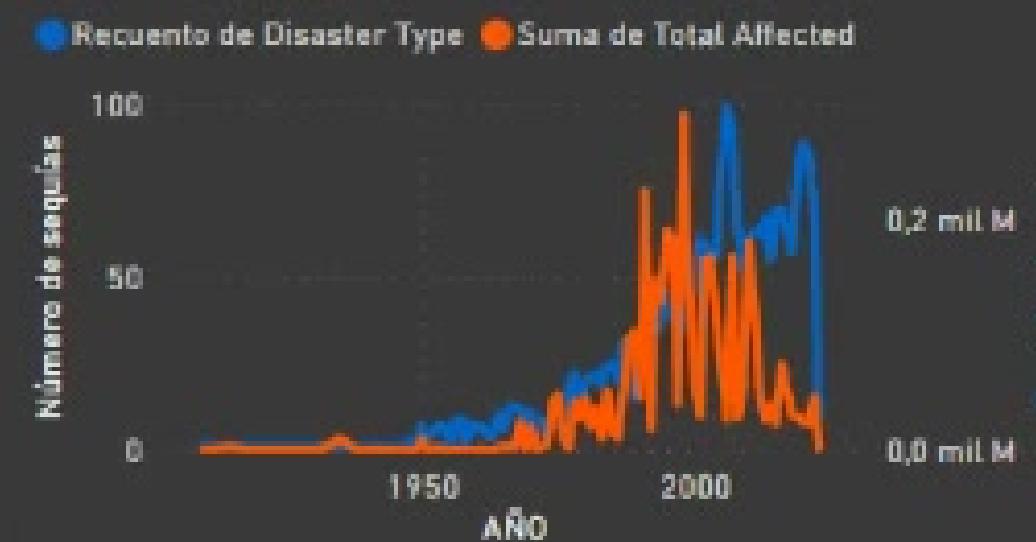
Frecuencia de inundaciones por continente



Suma de Total Afectados y Suma de Total de Muertes por País



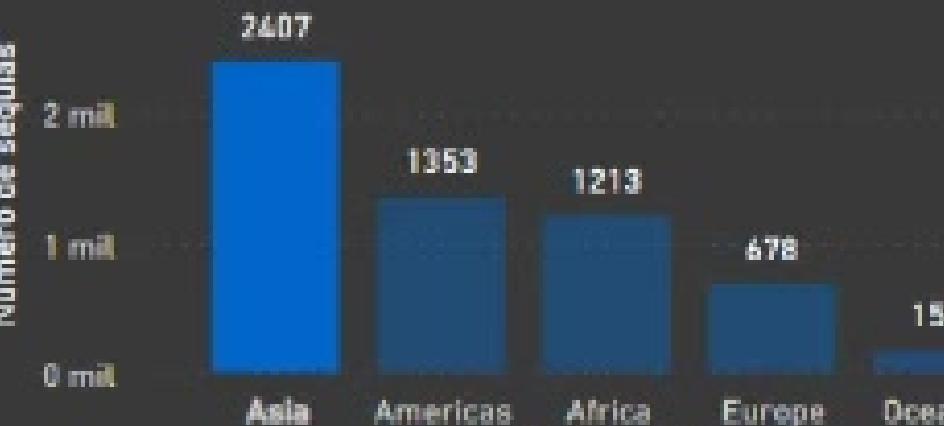
Análisis temporal por frecuencia y total de afectados



Fallecidos por tipo de desastre



Frecuencia de inundaciones por continente



Suma de Total Afectados y Suma de Total de Muertes por País



DESASTRES NATURALES

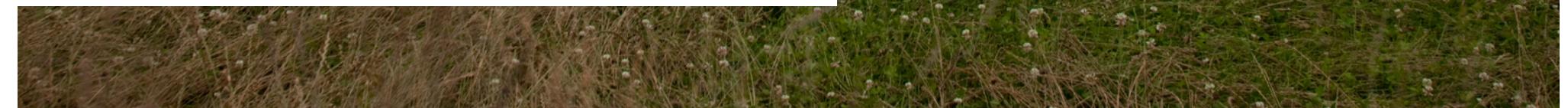
CONCLUSIONES

GENERALES

- Los DN se están incrementando con el tiempo.
- Reducción paulatina en pérdidas de vidas humanas, probablemente debido a medidas de protección poblacional.
- A partir de los 60 se evidencia tendencia a la alza en daños, que coincide con el aumento en la frecuencia de desastres.
- El análisis de datos es fundamental para organizar planes de emergencia frente a la repetición de desastres naturales.

SÍNTESIS DE IMPACTO:

MUERTES	AFFECTADOS	DAÑOS
DESCENSO	AUMENTO	AUMENTO



DESASTRES NATURALES

ESPECÍFICAS POR TIPO DE DESASTRE

Sequías:

- Las más letales.
- Incremento en sucesos y reducción del período entre los mismos.

Terremotos:

- Frecuencia irregular y a la alza hasta 2004.
- A partir de 70s, incremento de daños cuando magnitud supera 8.

Tormentas:

- Top 10 años con máximo record: 9 años son del SXXI.
- Picos de afectados cada 5 ó 6 años.

Inundaciones:

- Las más habituales.
- Más afectados.



MACHINE LEARNING

Modelo predictivo de Nivel de Destrucción



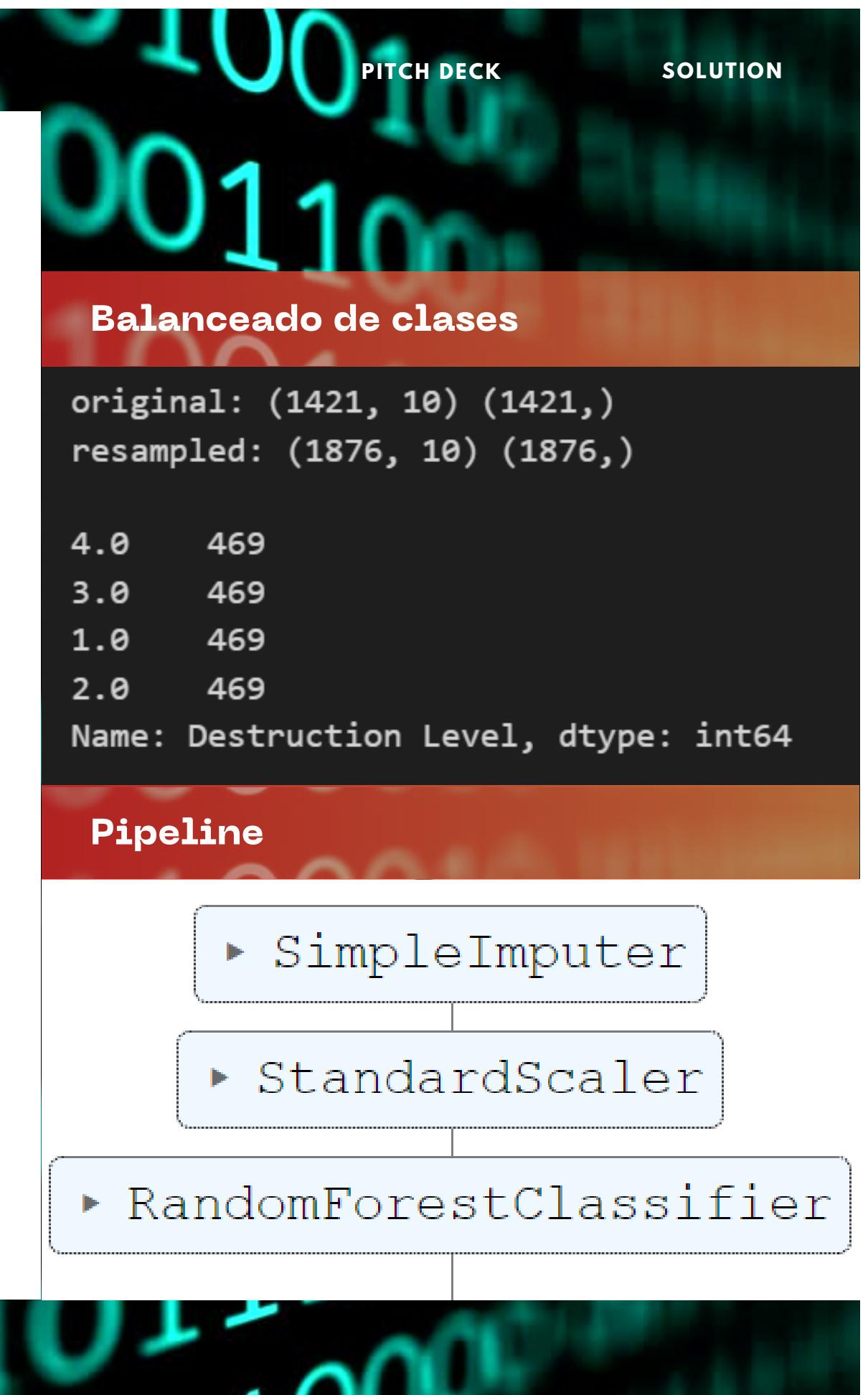
Herramientas



MODELO DE APRENDIZAJE SUPERVISADO

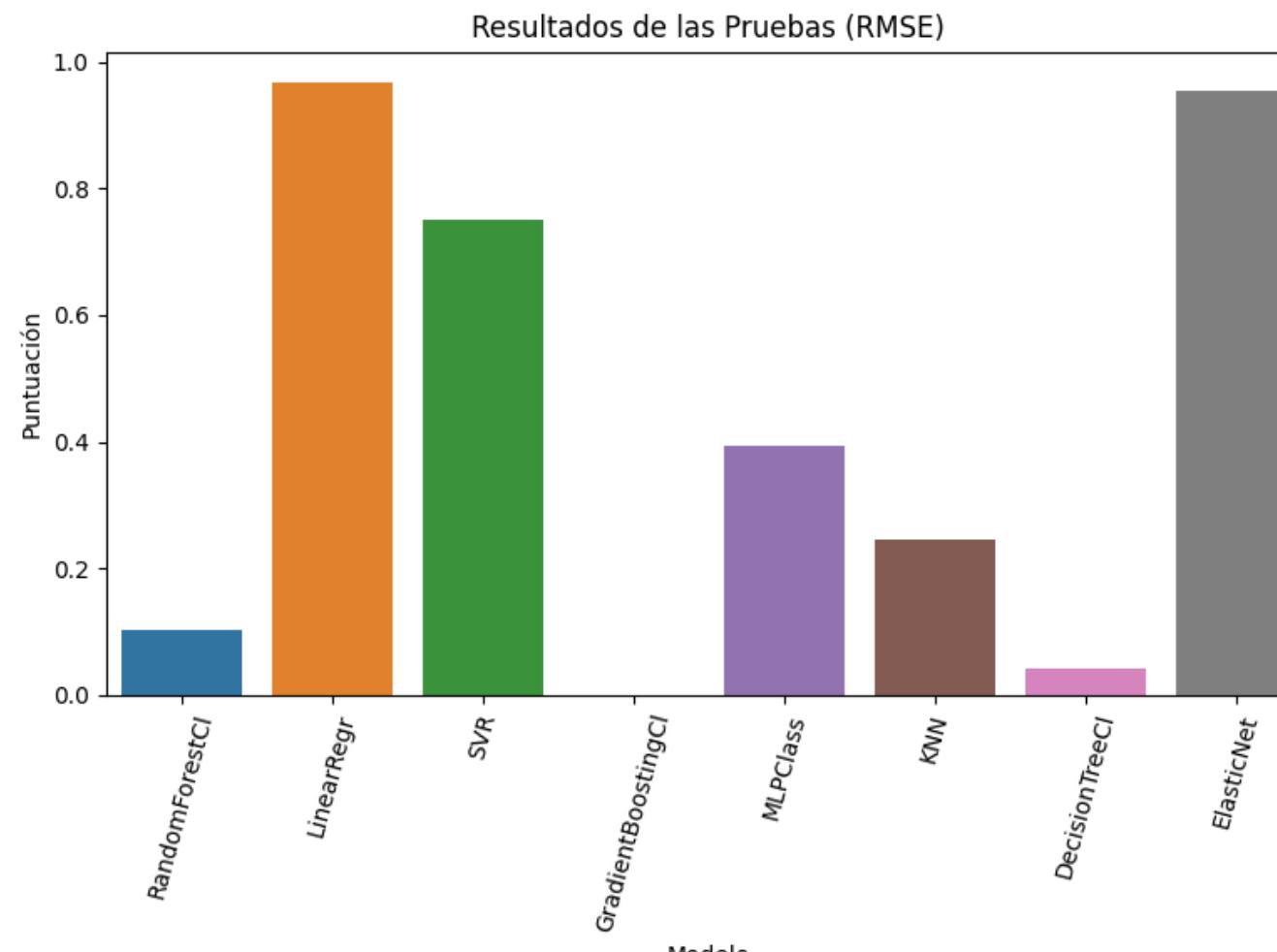
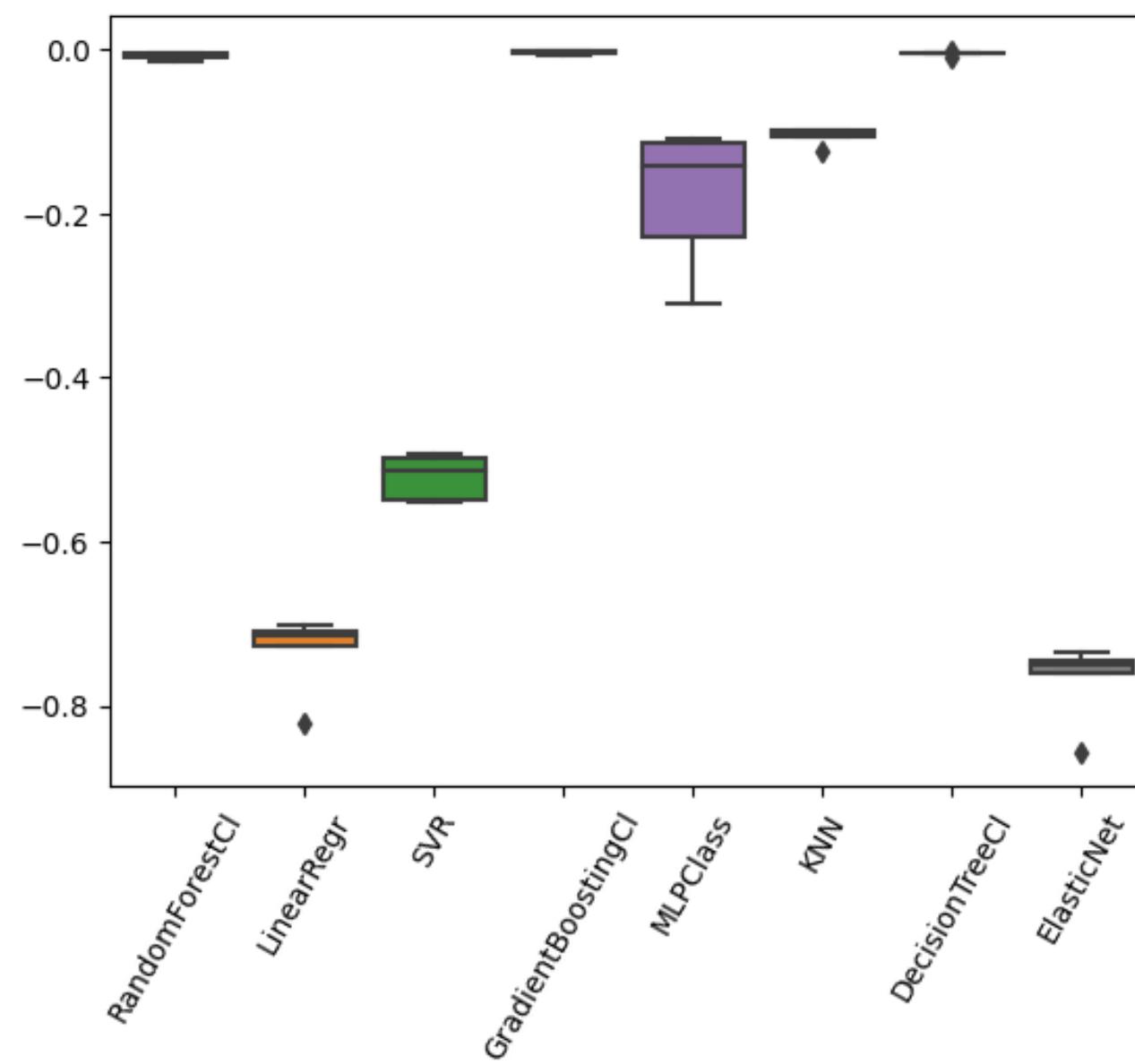
PROBLEMA DE CLASIFICACIÓN

- Basado en Valor de **Magnitud** de inundaciones (km²)
- Definición de clasificaciones del 'Low' 1, 'Moderate' 2, 'High' 3 y 4 'Very High' por **umbrales basados en cuartiles**.
- Función para asignar a cada clase de '**Nivel de Destrucción**' una combinación de **umbral y máximos de cada tipo de impacto**: fallecidos, afectados y daños.



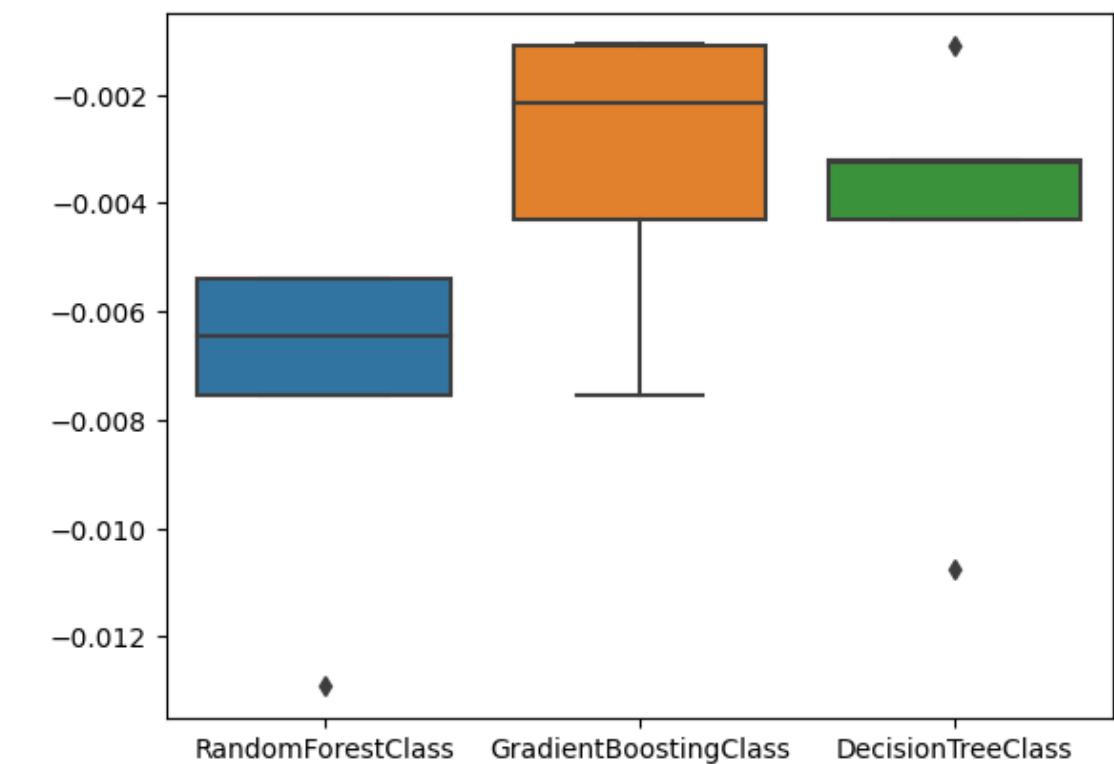
VALIDACIÓN DE ALGORITMOS

Análisis comparativo



Modelos finalistas:

RandomForestClassifier
GradientBoostingClassifier
DecisionTrees



PREDICCIÓN DE NIVEL DE DESTRUCCIÓN

Modelo ML con Random Forest Classifier

```
# Evaluar desempeño del dataset
print("Reporte de clasificación:")
print(classification_report(y_pred, y_pred))
```

Reporte de clasificación:

	precision	recall	f1-score	support
1.0	1.00	1.00	1.00	33
2.0	1.00	1.00	1.00	114
3.0	1.00	1.00	1.00	98
4.0	1.00	1.00	1.00	111
accuracy			1.00	356
macro avg	1.00	1.00	1.00	356
weighted avg	1.00	1.00	1.00	356

RESULTADOS DE CLASSIFICATION REPORT

Métricas colectivas de nivel de precisión y capacidad predictiva:

MUY ALTA

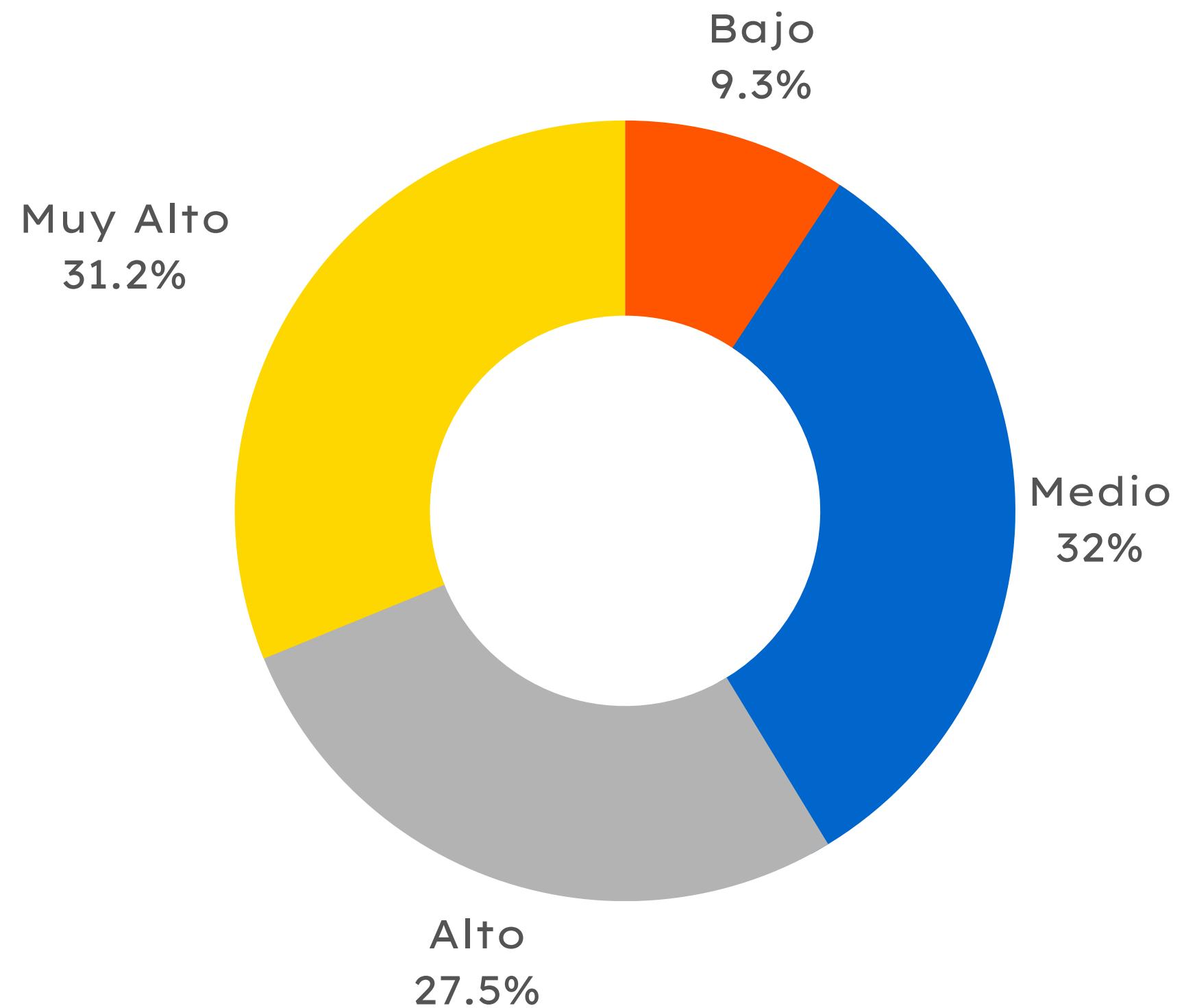
Iteración para mejora continua: tests y validación

Diagnóstico:

- 1) Posible overfitting.
- 2) Escasez de datos para niveles Bajo y Alto.
- 3) Entrenar con menos columnas.
- 4) Necesario ajustar umbrales con percentiles también en columnas de impacto.
- 5) Explorar otra configuración para umbrales no basada en estadística.

INUNDACIONES Y_PRED

PREDICCIÓN DEL NIVEL DE DESTRUCCIÓN



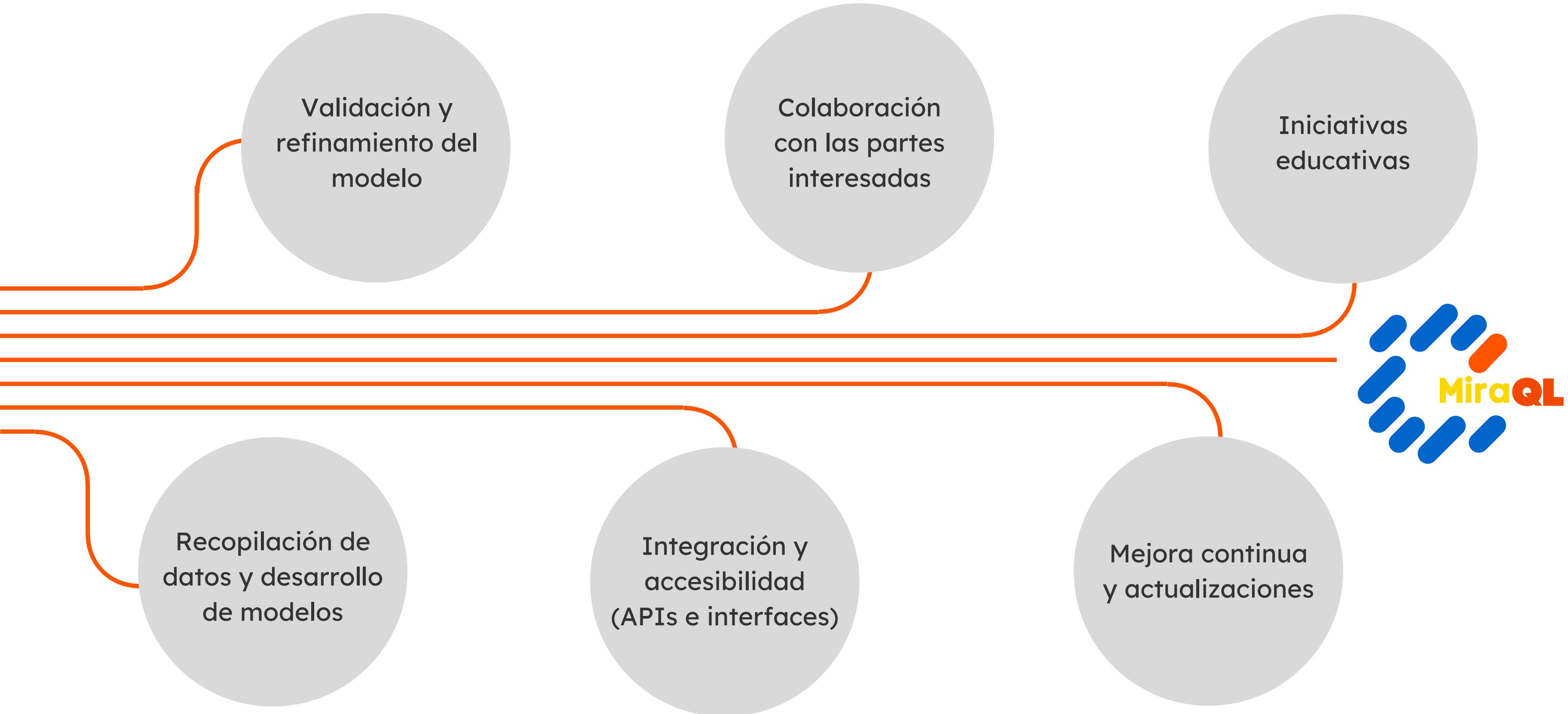
RECOMENDACIONES

Mejora de resultados de predicciones

- Más pruebas con datos nuevos de mayor volumen.
- Aplicar percentiles en distribución de umbrales de columnas de impacto.
- Preprocesado con menos etiquetas para X.
- Pipeline con modelos finalistas validados con desempeño alto como DecisionTrees y GradientBoostingClassifier.
- Nuevos ajustes a hiperparámetros.
- Iteración y reportes de clasificación de tests.

DESASTRES NATURALES CONCLUSIONES

IMPLEMENTACIÓN DEL MODELO PREDICTIVO

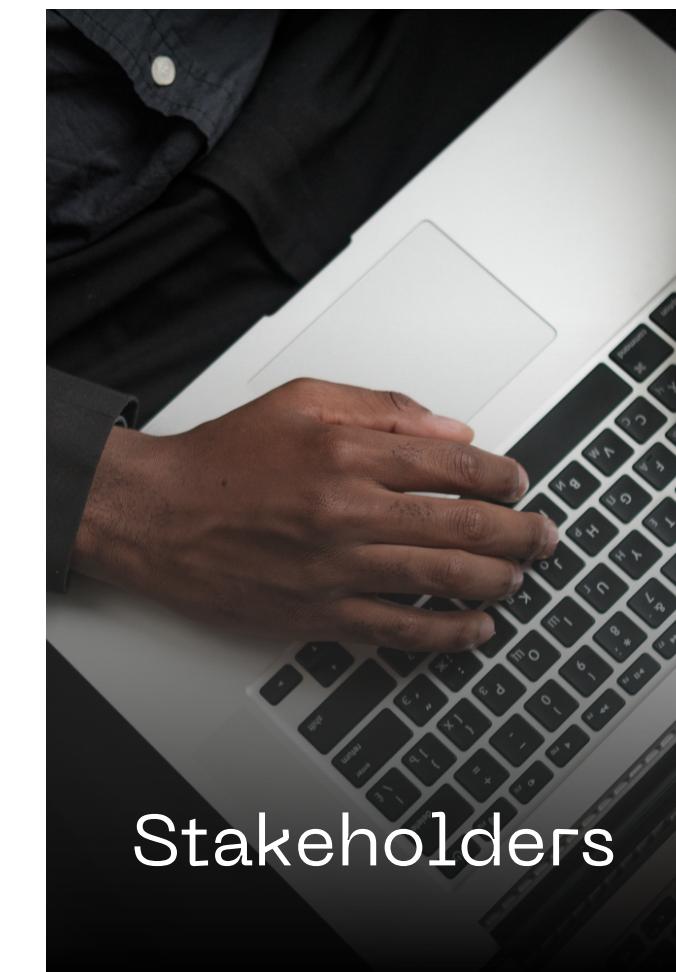


DESASTRES NATURALES
CONCLUSIONES

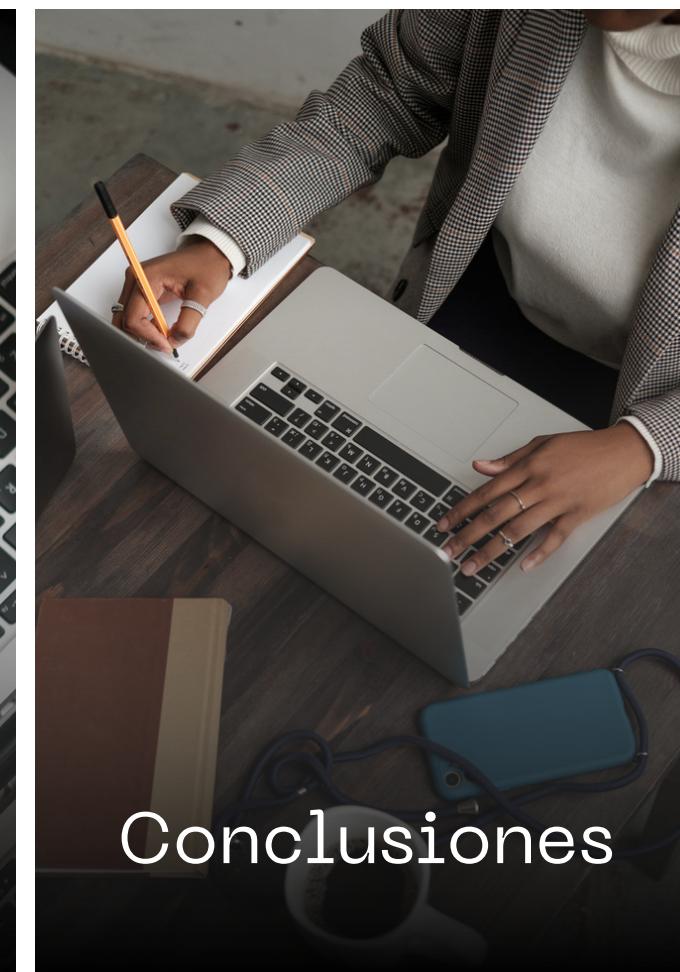
**MODELO DE
NEGOCIO,
CONCLUSIONES Y
RECOMENDACIONES**



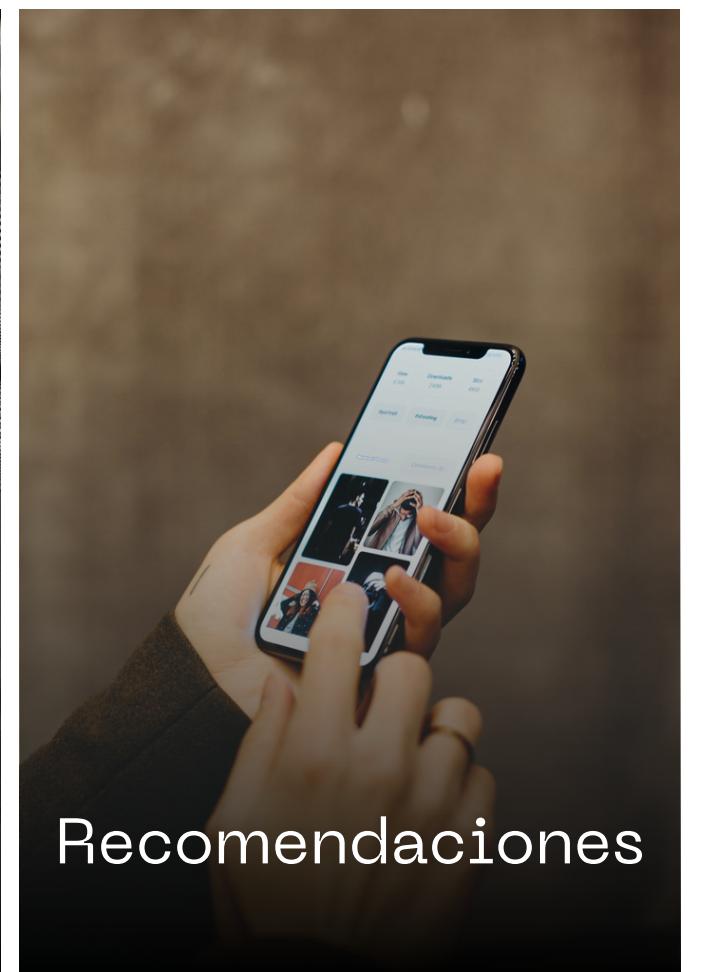
ANÁLISIS DE DATOS. EQUIPO 2



Stakeholders



Conclusiones



Recomendaciones

DESASTRES NATURALES

MODELO DE NEGOCIO EN CANVAS

KEY PARTNERS



Convenios de colaboración público-privada

- Administraciones públicas.
- Universidades.
- Agencias de recolección de datos sobre desastres.
- NASA y otras agencias aeroespaciales de observación metereológica satelital.
- Comunidades en riesgo.

KEY ACTIVITIES

- Consultoría BI.
- EDA.
- Diseño de Dashboards.
- CRISP-DM a BBDD sobre desastres naturales.
- Adaptación del modelo predictivo a diferentes tipos de desastres y regiones.
- Capacitación en uso de herramientas.

KEY RESOURCES

- BBDD de instituciones reconocidas con una metodología de recolección confiable.
- Patentes de modelos predictivos.
- Inteligencia de negocio.

COST STRUCTURE

- Hardware, licencias de software y actualización prof.
- Obtención de bases de datos de pago o por suscripción.
- Contratación de personal cualificado.
- Tecnología cloud para almacenamiento y gestión de grandes volúmenes de datos.
- Viáticos de desplazamiento para investigación en zonas en riesgo (en caso de que se requiera).
- Mejora continua de modelos y testing.

PROPIEDAD DE VALOR

Estrategia y planificación

Innovación

Gestión de ayuda y optimización del rescate

BI y especificidad sobre riesgos

Predicción es igual a prevención, protección y resiliencia

CUSTOMER RELATIONSHIPS

- In-company por consultoría y capacitación.
- Remotas para proyectos específicos como cualquier etapa de la metodología CRISP-DM EDA, dashboards, BI y ML.



CHANNELS

- Ferias y foros de tecnología de IA e inversión.
- In-company y remoto
- Blog, e-mail marketing y RRSS.
- Streaming para probar herramientas durante desastres naturales en tiempo real.



REVENUE STREAMS

Los clientes están dispuestos a pagar por:

- Tiempo es vida, no solo dinero.
- Minimización de desastres.
- Planificación de la reconstrucción en zonas de menor o cero impacto.
- Gestión y mejora continua de los modelos predictivos y la transformación de los datasets.
- Soluciones globales y regionales.

CUSTOMER SEGMENTS

- Administraciones públicas.
- Compañías de Seguros y Administradoras de Riesgos.
- Promoción Inmobiliaria.
- Cadena de Suministro y Logística.
- Desarrollo de Infraestructura y Planificación Urbana
- Servicios de Emergencia y Respuesta a Desastres.
- Comunidades y ONG.



*Partners pueden ser clientes, ya que aportan sus propios datos.





DIRIGIDO A:

**GOBIERNOS Y AGENCIAS
DE GESTIÓN DE
DESASTRES**



Estrategia

Mejorar estrategias de prevención, mitigación y respuesta.

**EMPRESAS
DE SEGUROS**



Especificidad sobre riesgos

Integrar en procesos de suscripción y evaluación de riesgos para ofrecer pólizas más precisas y ajustadas a riesgos específicos de cada región.

**EMPRESAS DE
CONSTRUCCIÓN E
INFRAESTRUCTURA**



Resiliencia

Identificar zonas de alto riesgo y diseñar edificaciones e infraestructuras resistentes.

**ORGANIZACIONES
DE AYUDA
HUMANITARIA**



Gestión de ayuda

Anticipar y preparar recursos y dirigir la ayuda hacia las áreas más afectadas.

**COMUNIDADES
VULNERABLES**



Protección

Preparación ante el riesgo y toma de decisiones informadas sobre la protección y evacuación.

DESASTRES NATURALES

PROCESO DE MEJORA CONTINUA

BENEFICIOS DE MEJORAS

- **Mayor Precisión**
- **Planificación Eficiente**
- **Respuesta Rápida**
- **Adaptabilidad a diferentes escenarios**



DESASTRES NATURALES

RECOMENDACIONES

GENERALES

- Consideraciones éticas sobre privacidad de datos y posibles sesgos.
- Combinar tecnología, experiencias multidisciplinarias y colaboración público-privada para abordar desafíos y estrategias de gestión de desastres.
- Integrar análisis de datos para infraestructuras y planes de emergencia.



DESASTRES NATURALES

RECOMENDACIONES

BASE DE DATOS:

- Incorporar magnitudes de densidad demográfica, densidad de construcción y/o duración.
- Recopilar BBDD específicas, normalizadas y homogéneas para cada tipo de desastre y modelo.
- Definir medida de magnitud para las sequias y su correlación con las personas afectadas.
- Reforzar y apoyar entes como EMDAT (1988), etc.



CONCLUSIÓN

- Concientizar sobre la dimensión del impacto de los desastres naturales
- La recurrencia de ciertos desastres naturales que afectan a determinadas poblaciones limitan seriamente las posibilidades de desarrollo de esas poblaciones.
- Impulsar el desarrollo de modelos predictivos de Desastres Naturales de grandes magnitudes.



GRACIAS

CONTACTA CON EL EQUIPO

Elena Blanco

blancoelenal1@gmail.com

LinkedIn:
[@elenablancodataanalytics](https://www.linkedin.com/in/elenablancodataanalytics)

Alejandra Eng

criptolmin@gmail.com

LinkedIn:
[@alejandraengbroca](https://www.linkedin.com/in/alejandraengbroca)

Rafael De Marco

rdemarcoz@outlook.com

LinkedIn:
[@rafael-de-marco-z-53097014](https://www.linkedin.com/in/rafael-de-marco-z-53097014)

Ángel Bardon

bardon.angel@gmail.com

LinkedIn:
[@angel-bardon-diez](https://www.linkedin.com/in/angel-bardon-diez)

AGRADECIMIENTOS



A todo el equipo, especialmente a Cristina, Pilar y Alan,
por habernos introducido al mundo del Análisis de Datos

A nuestros mentores:
Gema Piña de Minsait y José Luis Carvajal de INDRA

A las empresas que apoyan el programa de Escuela Digital
y a la Agencia para el Empleo del Ayuntamiento de Madrid