# Primer nivel. Caracterización de dominio

## Contextualización.

Chile es conocido como uno de los países con mayor actividad sísmica en el mundo. Su ubicación en el Cinturón de Fuego del Pacífico lo expone a frecuentes terremotos de gran magnitud. A lo largo de la historia, Chile ha experimentado numerosos terremotos significativos, incluyendo el terremoto de Valdivia en 1960, el más poderoso jamás registrado, con una magnitud de 9.5. Este fenómeno se debe a la subducción de la placa de Nazca bajo la placa Sudamericana, lo que genera una acumulación de tensión que eventualmente se libera en forma de sismos.

La alta frecuencia y magnitud de los terremotos en Chile hacen que el estudio y la visualización de estos datos sean de gran importancia para la comunidad. Para nosotros, como ciudadanos de Chile, es importante entender y estar en conocimiento del historial sísmico de nuestro país. Esta comprensión nos permite tomar mejores medidas de seguridad y desarrollar tecnología de prevención de daños. La preparación adecuada y la mejora continua en nuestras infraestructuras son fundamentales para enfrentar futuros terremotos con mayor resiliencia.

## Dataset utilizado

Para este objetivo, se utilizará el dataset chile\_earthquakes\_1520-2020.csv extraído de <https://www.kaggle.com/datasets/miguelciriano/chile-arthquakes-1705-2020>. Este dataset contiene datos de terremotos en chile desde el año 1520 hasta el año 2020 con datos extraídos del National Center for enviromental information. A través de 19 características, contiene información sobre la fecha exacta de ocurrencia de cada terremoto, región de ocurrencia, magnitud, profundidad, e información sobre las consecuencias tanto habitacionales, humanas y económicas. Cada ítem del dataset es un terremoto registrado.

Este dataset fue analizado y preprocesado en el archivo preprocesamiento.ipynb, en el que se trabajaron datos nulos, y se interpretaron algunos valores, para así después generar el archivo data.csv, el cual será utilizado para generar la visualización.

## Objetivos

De este modo los objetivos que busca este proyecto son:

1. Representar la ocurrencia histórica de los sismos en Chile.
2. Dejar en evidencia la evolución a través del tiempo, que han tenido los efectos de los terremotos en Chile. Buscar posibles factores correlacionados.
3. Concientizar sobre los efectos devastadores de estas consecuencias.

## Usuario objetivo

El potencial usuario es cualquier persona interesada en conocer los datos históricos sísmicos del país. También, es ideal para aquellos que desean comprender cómo ha evolucionado el impacto de los terremotos a medida que la sociedad chilena se ha desarrollado. Y por último esta herramienta es para aquellos que buscan concientizarse acerca de los peligros y el potencial de los terremotos.

Se espera que los usuarios tengan una experiencia fluida y amigable con la herramienta. En la que, basados en la simpleza y la claridad de las visualizaciones, se logren los objetivos propuestos. Esta interacción se realizará por medio de filtros de los datos, selección de la información específica a visualizar, y por último, búsqueda en detalle de ciertas características. No solo se mostrarán gráficos que presenten información concreta sobre alguno de los atributos del dataset sino también gráficos en los que 2 o mas atributos se encuentre relacionados, para identificar posibles correlaciones.

# Segundo nivel. Abstracción de datos y tareas

## Abstracción de datos

El dataset a trabajar, data.csv, posee 199 ítems (filas) y 19 atributos (columnas).

Se presentan solo los atributos que se utilizarán en la visualización.

|  |  |  |  |
| --- | --- | --- | --- |
| Atributo | Descripción | Abstracción | Tipo |
| Id | Identificador | Categórico secuencial | Llave |
| Year | Año de ocurrencia | Cuantitativo secuencial | Valor |
| Month | Mes de ocurrencia | Ordinal cíclico | Valor |
| Hour | Hora de ocurrencia | Cuantitativo cíclico | Valor |
| Tsunami\_associated | Hubo/No hubo tsunami asociado | Categórico | Valor (binario) |
| Region | Región de ocurrencia | Categórico | Valor |
| Latitude | Latitud del epicentro | Cuantitativo secuencial | Valor |
| Longitude | Longitud del epicentro | Cuantitativo secuencial | Valor |
| Focal\_depth | Profundidad del epicentro | Cuantitativo secuencial | Valor |
| Magnitude | Magnitud del terremoto | Cuantitativo secuencial | Valor |
| Deaths | Muertes provocadas | Cuantitativo secuencial (discreto) | Valor |
| Missing | Personas desaparecidas | Cuantitativo secuencial (discreto) | Valor |
| Injuries | Personas lesionadas | Cuantitativo secuencial (discreto) | Valor |
| Damage ($MM) | Daño estimado en millones de dolares | Cuantitativo secuencial | Valor |
| Houses\_destroyed | Casas destruidas por causa del terremoto | Cuantitativo secuencial (discreto) | Valor |
| Houses\_damaged | Casas dañadas por causa del terremoto | Cuantitativo secuencial (discreto) | Valor |

## Abstracción de tareas

1. Identificar correlación tiempo vs impacto. Identificar relaciones entre el tiempo y las distintas métricas del impacto de un sismo, dado un rango de magnitud.

Mediante tres distintos gráficos, el usuario podrá seleccionar un rango de magnitud y visualizar la relación entre el tiempo y las diferentes métricas de impacto. Las posibles métricas que cuantifican el impacto son: muertes, desaparecidos, lesionados, impacto económico, casas destruidas y casas afectadas.

El primer gráfico mostrará las métricas relacionadas al impacto sobre personas, el segundo el impacto económico y el tercero al impacto estructural, siempre sobre un rango de magnitudes, a lo largo del tiempo.

1. Descubrir las regiones con mayor actividad sísmica.

A través de un mapa geográfico de calor, se mostrarán las regiones de Chile indicando la actividad sísmica por intensidad de color. Además, el usuario podrá seleccionar una región específica y ver el detalle de la información.

1. Presentar los valores de profundidad focal de cada terremoto.

El usuario podrá activar el despliegue desde el mapa de las profundidades de cada terremoto. Mostrando una línea que se origina en su localización geográfica, y cuyo largo sea proporcional a la profundidad del foco. Cuando se seleccione una región específica los datos asociados a esta se resaltarán sobre el resto.

1. Preprocesar el dataset trabajando los datos faltantes. Conocer en que años y sobre que terremotos no se tienen registros de ciertos atributos.

## Tercer nivel. Codificación visual y de interacción

## Tarea 1

Para esta tarea se realizarán 3 gráficos que presenten los mismos elementos de diseño. Esto se debe a que son 3 gráficos que buscan representar un mismo objetivo, de identificar relaciones entre los efectos de cada terremoto con el paso del tiempo, únicamente variando la métrica que mide este efecto.

Así se decidió utilizar un gráfico de doble entrada en el que cada dato será representado por un punto, su posición vertical representará la magnitud del efecto en cuestión, mientras que su posición horizontal indicará la fecha de ocurrencia.

Todos los puntos de cada escala de medición serán unidos por una línea y serán representados con un color específico, para distinguir las escalas, pues varias escalas serán integradas en un mismo gráfico. La unión de los puntos por líneas se debe a que el objetivo es identificar relaciones, por tanto, queremos ver que tipos de recta se están dibujando en los gráficos y de este modo hacernos una idea de las relaciones existentes. El color solo será un atributo categórico, pues es para diferenciar las escalas.

También se podrá seleccionar un rango de magnitudes específico sobre los cuales se quiere visualizar los datos. Pues comparar los efectos de terremotos con respecto al tiempo, pero estos con distintas magnitudes es un escenario desigual y pierde significado. Lo interesante es ver como un terremoto de una misma escala afectaba antes y como está afectando ahora. Además, se podrá interactuar con lo gráficos mediante el paso del mouse sobre uno de los puntos, donde se desplegará el detalle específico del ítem que representa.

## Tarea 2

En esta tarea se representará un mapa de Chile. Se buscará mostrar la distribución geográfica de epicentros de los terremotos históricos a lo largo del país. La frecuencia con la que estos se produjeron se mostrará separándolos por regiones, y se utilizará una escala de intensidad de colores, formando un mapa de calor sobre el mapa geográfico. Esto se justifica porque podemos diferenciar fácilmente las regiones del país con mayor actividad sísmica solamente mirando el mapa.

Esta visualización está pensada para que interactúe con el usuario de manera que, si se presiona una región, esta queda seleccionada, desplegando el detalle de la cantidad total de sismos registrados en esta, y resaltándola visualmente por sobre las otras regiones del país. Además, esta misma selección sobre una región, filtrará los datos mostrados sobre los gráficos de la Tarea 1. Esto permitirá acotar los datos para ver los efectos de los terremotos específicamente sobre una región.

## Tarea 3

En esta tarea se utilizarán puntos sobre el mapa para representar cada terremoto. Deben ser marcas pequeñas, pues son muchos terremotos por registrar y tienen que caber todas las marcas dentro de un mismo mapa, sin saturar la visualización.

Desde cada una de estas marcas se desplegará una línea hacia abajo del mapa, que representará, según su altura, la profundidad del foco de cada terremoto. La línea será desplegada hacia abajo del mapa, pues hará mas intuitiva la representación de la profundidad, y al lado habrá un eje que indique la escala de profundidad. Los colores para cada una de estas marcas serán los mismos y no entregarán información.

Al seleccionar con el click una región específica, se oscurecerán todas las marcas de profundidad que no sean de la región seleccionada. Esto para fijarse en específico en cada región. Se podrá pasar el mouse por sobre cada marca y se desplegará el detalle del dato representado.

Como esta funcionalidad es una extensión de la gráfica del mapa de Chile, habrá un botón que permitirá desplegar o esconder esta funcionalidad.