

IMT 2200 - Introducción a Ciencia de Datos  
Pontificia Universidad Católica de Chile  
Instituto de Ingeniería Matemática y Computacional  
Semestre 2025-S2  
Profesor: Rodrigo A. Carrasco

## Proyecto de Investigación – Terremotos en Chile

Integrantes:

Darwin Zambrano

Nahuel Gutierrez

Jose Mora

## Motivación y contexto:

Dado que vivimos en uno de los países más sísmicos de América Latina y el mundo entero, consideramos indispensable el conocimiento sobre este fenómeno natural; lo que más atrae nuestro interés fue poder conocer las regiones en que mayor cantidad de terremotos se han producido, sabemos que existen métodos de advertencia sísmica pero queremos encontrar más métodos que puedan ayudar al país a prevenir estas catástrofes, esperamos que en base a los datos podamos encontrar un modelo que nos ayude a predecir los terremotos en los puntos más críticos de las regiones del país.

## Objetivo:

Con este proyecto esperamos poder identificar las regiones de Chile que más se han visto afectadas, así como también sus respuestas frente a estos sucesos, de modo que podamos reconocer los sectores que menos preparados se encuentran para un eventual sismo; las autoridades regionales se beneficiarían de esta información pues les ayuda a medir su preparación para posibles terremotos, la población también se beneficia por el conocimiento de esta información. Queremos calcular las zonas de mayor riesgo y observar cada cierto tiempo ocurren. Esperamos poder conseguir un método que pueda ayudar a predecir la eventual ocurrencia de estas catástrofes.

## Datos:

Los datos serán extraídos de una fuente pública, de la organización USGS Scientific Research Results, en este link pueden encontrar la página donde se extrae la información.  
<https://www.usgs.gov/programs/earthquake-hazards/earthquakes> . La información obtenida es en formato .csv, que fue descargada mediante el siguiente link: <https://earthquake.usgs.gov/earthquakes/search/>, dicho documento descargado viene en un formato estructurado, tiene un peso de 1.7mb, y al convertirlo en Data Frame tiene más de 10.000 columnas y 22 filas con información como: tiempo, latitud, longitud, ubicación, magnitud, escala, entre otros. Más información la pueden encontrar en nuestro repositorio de Github en el siguiente link:  
<https://github.com/Skysilen-t/Proyecto-Ciencia-de-Datos/tree/main/Propuesta%20y%20documento>

## Preguntas:

- Obtener información de los terremotos en Chile  
(¿Qué información es posible obtener respecto de los terremotos en Chile?)
- Calcular los mayores terremotos por región  
(¿Cuáles han sido los mayores terremotos de cada región?)
- Calcular las zonas de mayores riesgos para terremotos  
(¿En qué zonas hay mayor riesgo de terremotos?)
- Crear un modelo que pueda predecir en base a las estadísticas los posibles lugares de riesgos  
(¿Podemos saber cuales son los lugares con más riesgo de terremotos?)
- Crear un modelo estadístico que puede predecir cada cierto tiempo pueden acontecer un terremoto  
(¿Cada cuanto puede ocurrir un nuevo terremoto?)

## Diseño tentativo:

Primero queremos limpiar la información, que no tenga datos nulos, filtrar solo por los terremotos ocurridos en Chile, luego crear un Data Frame por cada región en Chile, para filtrar y encontrar en que comunas ocurrieron la mayor cantidad. Una vez ubicados estos sitios haremos una gráfica con respecto al tiempo en que ocurren los terremotos en dicha ubicación. En base a esta gráficas queremos observar si siguen algún patrón que nos ayude a predecir futuras catástrofes.