#### Análisis de Tendencias de Defunciones en Chile (1990-2021)

La presente propuesta de proyecto, tiene como objetivo relacionar varios datasets: ingresos hospitalarios y las defunciones en Chile, entre los años 1990-2021. Esta motivación surge de problemáticas interesantes, como los desafíos que tiene el sistema de salud, relacionado a enfermedades crónicas más comunes, accidentes, envejecimiento de la población, disparidades en los accesos a la salud por región, entre otros.

También el análisis busca responder preguntas valiosas para la prevención y conciencia en la población, además del mismo sistema de salud, para poder adoptar medidas claves en la gestión sanitaria y en la atención hospitalaria. Con la fusión de los datasets es posible analizar las saturaciones de los hospitales o las causas de mortalidad más comunes en Chile. Además es importante revisar la calidad de los datos proporcionados por el Ministerio de Salud (MINSAL). Hurgar en la precisión y veracidad de estos datos tiene un aspecto crucial, ya que, estos errores podrían afectar directamente a la población. Por lo anterior, esta información puede ser de ayuda para gestores de salud, la población común, y autoridades sanitarias.

#### Descripción de los Datos:

Ambos datasets han sido recopilados desde la fuente pública de datos del departamento de estadísticas e información de salud del gobierno, obtenido de: <a href="https://deis.minsal.cl/">https://deis.minsal.cl/</a>.

#### 1) Defunciones por causa 1990-2021:

- Descripcion de cada variable y tipo:
- https://github.com/vaniaulloa/ATDC--imt2200/tree/main/defuncione s%201994-2021
- formato: archivo csvvolumen: 04 Mb

### 2) Egresos hospitalarios:

 Descripcion de cada variable y tipo: <a href="https://github.com/vaniaulloa/ATDC--imt2200/tree/main/Datos%20e">https://github.com/vaniaulloa/ATDC--imt2200/tree/main/Datos%20e</a> gresos%20hospitalarios

- formato: archivo csv

- volumen: 5,75 GB (datos desde el 2001 hasta el 2021)

- Preguntas de investigación: proponga al menos 5 preguntas específicas que buscarán responder con base en los datos presentados. Estas preguntas pueden ser relativas a relaciones entre variables, capacidad de predicción de una variable, clasificación, identificación de clusters o categorías, etc.
  - 1) ¿Cuál es la causa externa más común dentro de los datos analizados? se considera causa externa, como aquellas no relacionadas a enfermedades.
  - 2) ¿Cómo influyen los periodos de saturación hospitalaria en la tasa de mortalidad en ciertas patologías?
  - 3) ¿Existe una tendencia con las causas de muerte en cada uno de los grupos etarios? considerando también las disparidades de género en las tasas de mortalidad.
  - 4) ¿Qué tan efectivas son las hospitalizaciones en la reducción de la mortalidad por enfermedades? Aquí consideramos una investigación, con las causas más comunes de muerte por enfermedades, considerando una diferencia entre enfermedades tratables y no tratables.
  - 5) ¿Existen factores que explican las variaciones en las tasas de mortalidad y hospitalizaciones? Considerando un análisis por región (demográfico, socioeconómico, otros).

Para este análisis, utilizaremos varios métodos. Dentro de los **métodos computacionales**, tenemos planeado emplear los siguientes:

- Análisis Exploratorio de Datos (EDA): Visualización de datos a través de gráficos como histogramas, diagramas de dispersión y mapas de cloropletas, entre otros que se consideren adecuados según cada ítem y/o pregunta, lo que nos permitirá analizar tendencias y patrones relacionados con nuestras preguntas de investigación.
- Preprocesamiento de Datos: Limpieza y transformación de los datos, que incluye la eliminación de duplicados, el manejo de valores nulos, la normalización y la fusión de nuestras base de datos. Esto garantizará la calidad de los datos para su análisis.
- Fusión de Datos: Integraremos diferentes fuentes de datos para analizarlas en conjunto y responder varias de nuestras preguntas, lo que nos permitirá extraer conclusiones importantes y descubrir relaciones entre los conjuntos de datos.

En cuanto a los **métodos estadísticos**, planeamos usar los siguientes:

- Medidas de tendencia central: como la media, mediana y moda, para describir la distribución de los datos, en diferentes ciudades o en general.
- Tasa de Mortalidad General (TMG): Indicará el número de defunciones por cada 1.000 habitantes en un año determinado.
- Tasa de Mortalidad Específica por Causa (TMEC): Refleja el número de defunciones por una causa específica por cada 100,000 habitantes, permitiendo un análisis más detallado de los factores que influyen en la mortalidad.

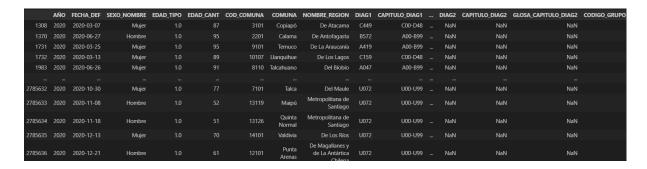
La muestra de los datos, se puede obtener mediante el repositorio del proyecto, que se encuentra en el siguiente enlace:

https://github.com/vaniaulloa/ATDC--imt2200/tree/main/Muestras

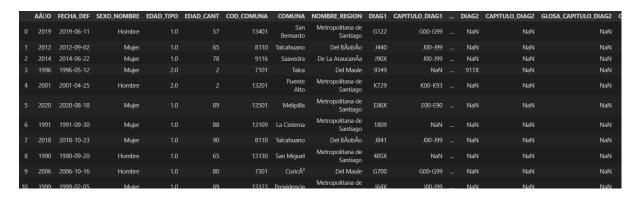
Estas se encuentran como archivos del formato csv, cada una contiene 20 datos, seleccionados aleatoriamente, a modo ilustrativo.

#### Visualización de los dataframes:

- Defunciones:



## Egresos hospitalarios:



# Bibliografía

- Minsal. <a href="https://deis.minsal.cl/">https://deis.minsal.cl/</a> (sección datos abiertos)