

# Informe: Principales Causas de Muerte en Argentina

## 1. Descripción del Dataset y Justificación

El dataset seleccionado para este análisis es una combinación de dos conjuntos de datos oficiales: "Defunciones ocurridas y registradas en la República Argentina entre los años 2005-2022" y "Defunciones ocurridas y registradas en la República Argentina 2023", ambos provenientes del Ministerio de Salud de Argentina, específicamente de la Dirección de Estadística e Información en Salud (DEIS). Este dataset recopila información sobre las defunciones ocurridas y registradas en el país durante el período de 2005 a 2023. Incluye variables sociodemográficas clave como la provincia de residencia, sexo, causa de la defunción, edad al momento del fallecimiento y una especificación sobre si la muerte fue materna.

Se eligió este dataset por varias razones, en primer lugar, su amplio rango temporal de casi dos décadas permite realizar un análisis profundo, identificando tendencias y patrones en las tasas de mortalidad a lo largo del tiempo. En segundo lugar, la riqueza de sus variables sociodemográficas posibilita comparaciones entre diferentes grupos poblacionales, lo que es fundamental para comprender disparidades en los resultados de salud. En conjunto, estas cualidades hacen de este dataset una base sólida y versátil para explorar diversos aspectos de la mortalidad en Argentina, ideal para esta presentación de análisis de datos.

## 2. Objetivo de análisis

El objetivo principal de este análisis es:

- Analizar la evolución anual de la cantidad de muertes asociadas a las tres causas más frecuentes de defunción en el país.
- Comparar la distribución porcentual de estas causas en función de los distintos grupos etarios, para identificar qué segmentos de la población se ven más afectados en términos relativos.
- Visualizar patrones y tendencias a lo largo del tiempo que permitan interpretar cambios en el perfil epidemiológico de la población argentina.

Este enfoque permite comprender el comportamiento de las principales causas de defunción a través del tiempo y en diferentes etapas de la vida, aportando así una mirada integral desde la salud pública.

## 3. Principales hallazgos / insights

### Principales Causas de Defunción

Las tres principales causas de defunción identificadas en Argentina durante el período analizado son:

- Insuficiencia cardíaca
- Neumonía, organismo no especificado
- Infarto agudo del miocardio

Estas causas representan más del 20% del total de defunciones registradas y son fundamentales para entender las tendencias de salud en el país.

### Tendencias Temporales

El análisis temporal ha revelado los siguientes patrones:

- Insuficiencia cardíaca:** Fue la causa predominante de defunción desde el año 2005 hasta 2014.
- Neumonía, organismo no especificado:** Se convirtió en la principal causa de defunción a partir del año 2015 hasta 2023.
- Infarto agudo del miocardio:** Esta causa se ha mantenido constante a lo largo del tiempo, con una ligera tendencia ascendente.

Además, se observaron picos de incidencia específicos para cada causa:

- La Insuficiencia cardíaca alcanzó su punto máximo de defunciones entre los años 2010 y 2011.
- La Neumonía, organismo no especificado tuvo su pico en el año 2022.
- El Infarto agudo del miocardio tuvo su pico entre los años 2021 y 2022

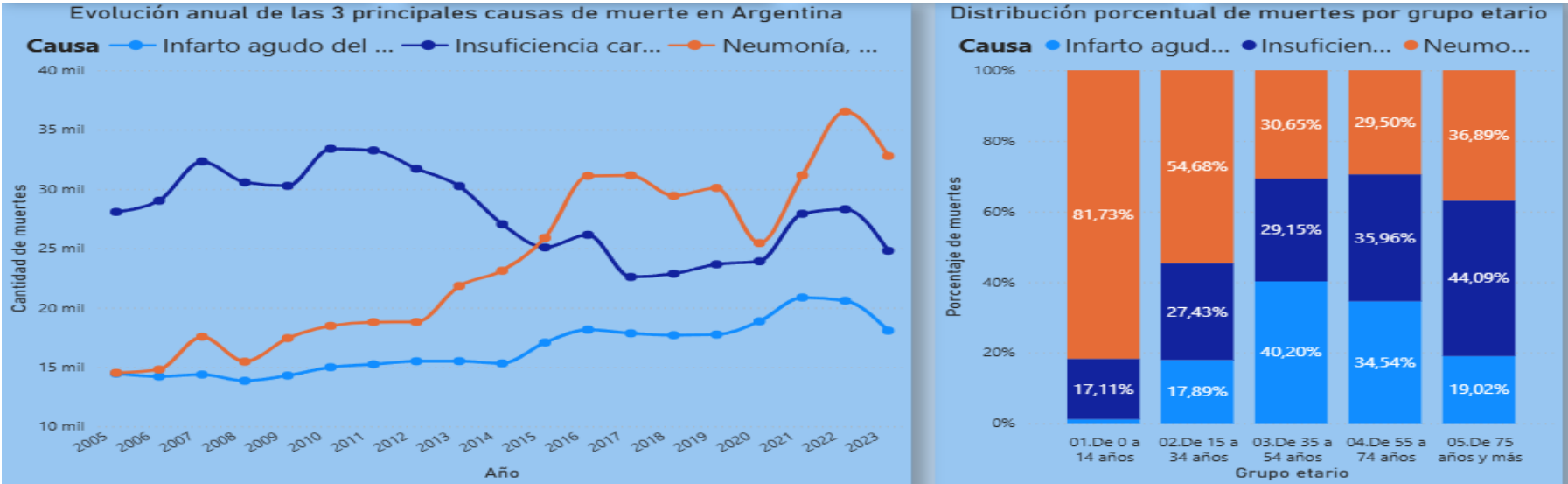
### Distribución por Grupo Etario

Un análisis más detallado de la distribución por grupo etario ha revelado destacables diferencias en las causas de defunción más frecuentes:

- De 0 a 14 años:** La principal causa de defunción es la Neumonía, organismo no especificado.
- De 15 a 34 años:** Igualmente, Neumonía, organismo no especificado es la causa más frecuente.
- De 35 a 54 años:** El Infarto agudo del miocardio se presenta como la principal causa de defunción.
- De 55 a 74 años:** La Insuficiencia cardíaca es la causa predominante.
- De 75 años y más:** La Insuficiencia cardíaca sigue siendo la principal causa de defunción.

Estos hallazgos subrayan la importancia de considerar la edad al diseñar intervenciones de salud pública, ya que las causas de defunción varían significativamente entre los diferentes grupos etarios.

## 4. Breve descripción de las visualizaciones



### Visualización 1: Evolución anual de las 3 principales causas de muerte en Argentina

Este gráfico de líneas representa la cantidad absoluta de muertes anuales para cada una de las tres causas principales identificadas: Infarto agudo del miocardio (color celeste), Insuficiencia cardíaca (color azul) y Neumonía, organismo no especificado (color naranja).

- **Eje X:** Muestra el rango de años, desde 2005 hasta 2023, permitiendo observar la trayectoria temporal de cada causa.
- **Eje Y:** Indica la cantidad de muertes en miles, proporcionando una escala clara para comparar la magnitud de cada causa a lo largo del tiempo.

A través de este gráfico, se puede observar visualmente el predominio inicial de la "Insuficiencia cardíaca", el posterior ascenso y liderazgo de la "Neumonía, organismo no especificado", y los picos de cada una de estas causas en los años mencionados en la sección de hallazgos.

### Visualización 2: Distribución porcentual de muertes por grupo etario

Este gráfico de barras apiladas al 100% usa los mismos colores que el anterior y muestra la composición porcentual de las tres principales causas de defunción dentro de cada uno de los grupos etarios definidos.

- **Eje X:** Representa los cinco grupos etarios clasificados: "01. De 0 a 14 años", "02. De 15 a 34 años", "03. De 35 a 54 años", "04. De 55 a 74 años" y "05. De 75 años y más".
- **Eje Y:** Indica el porcentaje del total de muertes dentro de cada grupo etario que corresponde a cada una de las tres causas principales, sumando un 100% en cada barra.

El tamaño de cada segmento es proporcional a su porcentaje en ese grupo de edad. Esta visualización facilita la identificación inmediata de qué causa de defunción es la más prevalente en cada etapa de la vida, así como la proporción que cada una de las tres principales causas aporta a la mortalidad total dentro de ese rango etario.

Ambas visualizaciones son complementarias y proporcionan una comprensión clara y concisa de los patrones de mortalidad en Argentina, tanto en su evolución histórica como en su impacto diferenciado por edad.

## 5. Reflexión sobre el proceso

### Dificultades Encontradas:

Una de las principales dificultades fue la gestión de la calidad y completitud de los datos. El dataset, que abarca un período extenso (2005-2023), presentaba una gran cantidad de datos nulos, especialmente en las columnas "muerte\_materna\_id" y "muerte\_materna", donde todas las defunciones no maternas se registraron con campos vacíos, también las columnas "jurisdiccion\_residencia\_nombre" y "cie10\_clasificacion" presentaban una pequeña cantidad de campos vacíos.

### Decisiones Clave:

1. **Herramientas:**
  - Se optó por utilizar Python, con la librería Pandas, para la combinación de los dos conjuntos de datos.
2. **Tratamiento y depuración de datos:**
  - Se decidió eliminar todos los datos duplicados, mientras que los valores nulos fueron reemplazados por valores específicos según cada columna. En el caso de la columna "cie10\_clasificacion", se consultó la documentación de CIE-10 y se emparejó cada código de enfermedad con su identificador en "cie10\_causa\_id", generando así una nueva versión de "cie10\_clasificacion" sin campos vacíos.
3. **Selección de variables:**
  - Se decidió centrarse en las tres principales causas de defunción debido a su impacto significativo en la mortalidad y su relevancia para la salud pública.
4. **Agrupación por Rangos Etarios:**
  - Se excluyó del análisis por grupo etario la categoría "06.Sin especificar", para poder centrar el análisis en rangos etarios definidos y comparables.

En resumen, aunque el proyecto presentó desafíos, no solo pude generar hallazgos concretos sobre la mortalidad en Argentina, sino que también me permitió desarrollar una apreciación más profunda por la complejidad del análisis de datos en el contexto de la salud pública.