# FICHA TÉCNICA DE BASE DE DATOS DE TASAS DE DEFUNCIONES POR CANCER

## Objetivo general del proyecto.

Diseñar y crear una plataforma para coordinar esfuerzos de intercambio seguro de datos de salud y la creación de servicios seguros de analítica de datos sin realizar cambios mayores en los sistemas de expediente Clínico Electrónico (SECE) existentes permitiendo su adecuación al cumplimiento de normas oficiales y considerando los estándares que rigen su implementación. La plataforma se basará en servicios configurables para la gestión, aseguramiento, control de acceso, intercambio y preservación de grandes volúmenes de datos en salud (Big Data) enfocados en avanzar hacia la creación de un SECE único nacional y coadyuvar en la edificación de un repositorio nacional de ejecución segura de servicios de análisis de datos médicos.

## Propósito del recurso

Este conjunto de datos es una recopilación de las diferentes bases de datos que se encuentran en DGIS en un solo archivo para su uso en aplicaciones de big data y GIS (Sistemas de información georreferenciada).

Propósito: Esta base de datos se utilizó para probar sistemas de procesamiento de datos de mortalidad para realizar estudios prospectivos y exploratorios de la mortalidad por cáncer en el territorio Nacional, se transfirió al Registro Nacional de Cáncer con Base poblacional de IMSS-INCAN, en donde actualmente es utilizada para realizar búsquedas de datos.

Diferencia con la base original:

La diferencia entre los datos de DGIS y la base de cáncer creada por la plataforma Xelhua son los siguientes:

1. Reducción: La base de “Tasas Ajustadas de defunciones por cáncer por municipio del 2000 al 2020” se basa en una versión depurada, curada y preprocesada de la creada por la DGIS, la cual fue filtrada para incluir solo casos de defunción por Cáncer.

2. Procesamiento: La base filtrada/reducida fue procesada para obtener tasas crudas por estados/municipio y ajustadas a los rangos poblacionales definidas por el RNC-BP y sugeridas por la OMS. La base también fue procesada para calcular las posiciones de georreferenciación para cada municipio y estado.

3. Enriquecimiento: La base fue enriquecida con variables categóricas: tasas crudas y ajustadas, así como las posiciones de georreferenciación.

En resumen: esta base cuenta con el cálculo de tasas ajustadas con base a la población de cada municipio, según la población de cada sexo, rango de edad y el año correspondiente. Además, según el volumen de la población se calcularon diferentes tasas: por cada 100 mil habitantes (100k), 10k y 1k habitantes, lo cual fue calculado dependiendo de la población total de cada municipio.

## Investigador principal a cargo.

Juan Armando Barrón Lugo

## Recolección/obtención de los datos.

La recolección de estos datos fue mediante la plataforma de datos abiertos de la secretaria de Salud de la República Mexicana (<http://www.dgis.salud.gob.mx/contenidos/basesdedatos/bdc_defunciones_gobmx.html>), estos datos fueron descargador por año y filtrados por el código CIE-10 con inicial C, correspondiente a los diferentes tipos de cáncer.

## Periodo de recolección/obtención de datos.

El periodo de recolección de los datos es de 2000 al 2020, utilizados dentro del proyecto. Actualmente se encuentra el año 2021 disponible en los datos abiertos.

## Variables Incluidas

Cuenta con un total de 12 variables descritas a continuación:

* **ANIO\_REGIS:** Año de registro de la defunción.
* **ENT\_CVE:** Identificador de la entidad federativa de ocurrencia.
* **MUN\_CVE:** Identificador del municipio de ocurrencia.
* **CAUSA\_DEF:** Identificación de la causa de defunción. Dado que este es un subconjunto del conjunto de datos de defunciones en general todos los identificadores corresponden a cáncer comenzando por la letra ‘C’ seguida de 2 dígitos.
* **SEXO:** Identificador del sexo del difunto.
* **RANGO\_EDAD:** Rango de edad del difunto.
* **Descrip:** Descripción del tipo de causa de defunción. Dado que este es un subconjunto del conjunto de datos de defunciones en general todos los valores corresponden a subtipos de cáncer (359 subtipos únicos).
* **CONTEO:** Conteo de defunciones por tipo de cáncer, sexo, municipio y rango de edad determinado.
* **POBLACION:** Población según el municipio, sexo y rango de edad determinado.
* **nombre municipio:** Nombre del municipio de ocurrencia.
* **TASA:** Tasa de defunción ajustada a la población
* **TASA\_TYPE:** Tipo de tasa calculada con respecto al volumen de la población (cada 100, 1k, 10k, y 100k habitantes si la población es menor a 100, mayor a 100 y menor a 1k, mayor a 1k y menor a 10k, y mayor a 100k habitantes respectivamente)

## Estrategia de aseguramiento para la proyección de datos sensibles/personales.

No aplica debido a que los registros son abiertos y no cuentan con datos personales.

## Fecha última de actualización.

Ultima fecha de actualización es el viernes 16 de diciembre de 2022.

## Mencionar el tipo de **controles para la validación y verificación** de la captura/registro de los datos.

No aplica.

## Otras **plataformas** donde se encuentren disponibles estos recursos de información. Incluir URL.

Carpeta compartida:

<https://1drv.ms/u/s!AtMgnjYpElvzg5dmjBS-3RFcSb9i4Q?e=5ynUIh>

Tablero de consulta DGIS.

<http://sinaiscap.salud.gob.mx:8080/DGIS/>

## Otras fuentes de Financiamiento.

No aplica.

## Seguimiento de la Cohorte en estudio.

No aplica.

## Publicaciones.

No aplica.

## Otros. Cualquier información extra que sea pertinente precisar o agregar.