



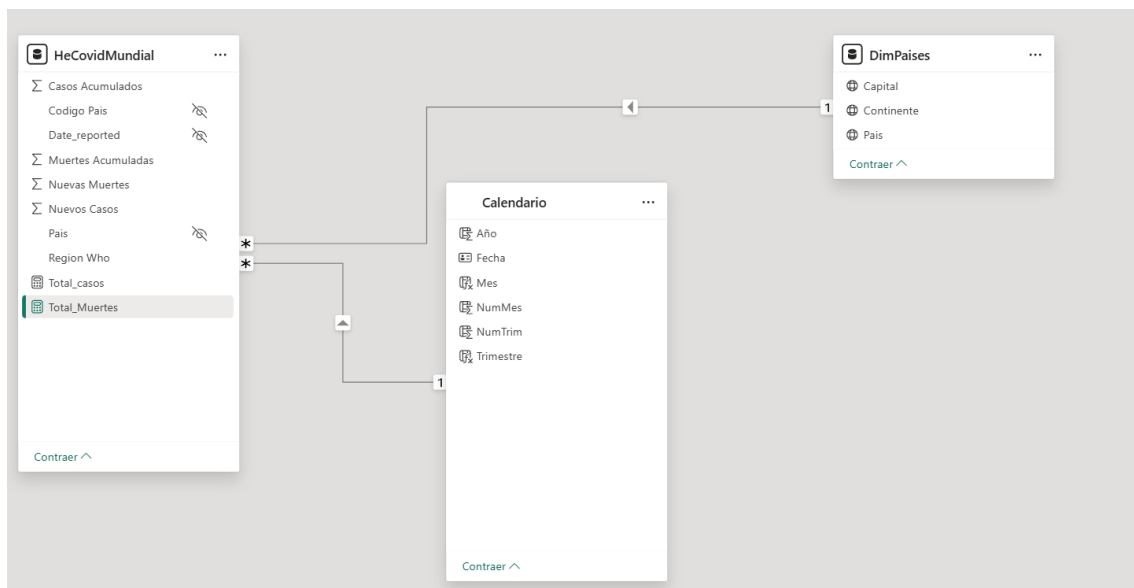
## TAREA 3 – Módulo 2

### Asignatura: Big Data

**Nombre:** Rodrigo Mora

**Fecha:** 09/08/2025

#### 5.1. Capturas de pantalla del modelo copo de nieve resultante estableciendo una descripción.



En la imagen se muestra un modelo de datos en esquema copo de nieve desarrollado en Power BI, el cual está compuesto por una tabla de hechos y dos tablas de dimensiones. Este modelo nos permite organizar y analizar la información de los casos y muertes por COVID-19 de forma más eficiente.

#### Tabla de hechos: HeCovidMundial

Aquí se guardan los datos numéricos más importantes como:

- Casos Acumulados y Nuevos Casos.
- Muertes Acumuladas y Nuevas Muertes.
- Medidas calculadas como Total\_casos y Total\_Muertes.
- Además, contiene campos clave como Codigo Pais, Date\_reported y Pais.
- Esta tabla se relaciona:
- Con la tabla Calendario por el campo de fecha (Date\_reported ↔ Fecha), en una relación de muchos a uno.
- Con la tabla DimPaises por el campo Pais, también en una relación de muchos a uno.

### **Tabla de dimensión de tiempo: Calendario**

Contiene datos que describen el tiempo, como Año, Mes, Número de Mes (NumMes), Trimestre y Número de Trimestre (NumTrim).

Tiene una relación de uno a muchos con HeCovidMundial, ya que cada fecha del calendario puede aparecer varias veces en los registros de casos.

### **Tabla de dimensión geográfica: DimPaises**

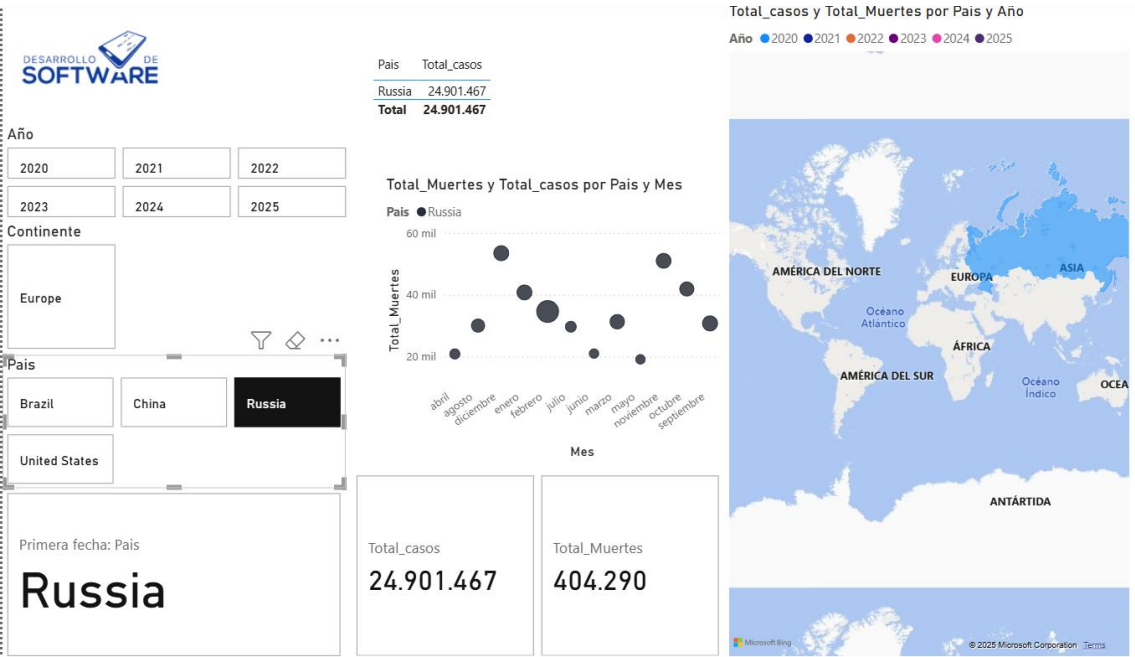
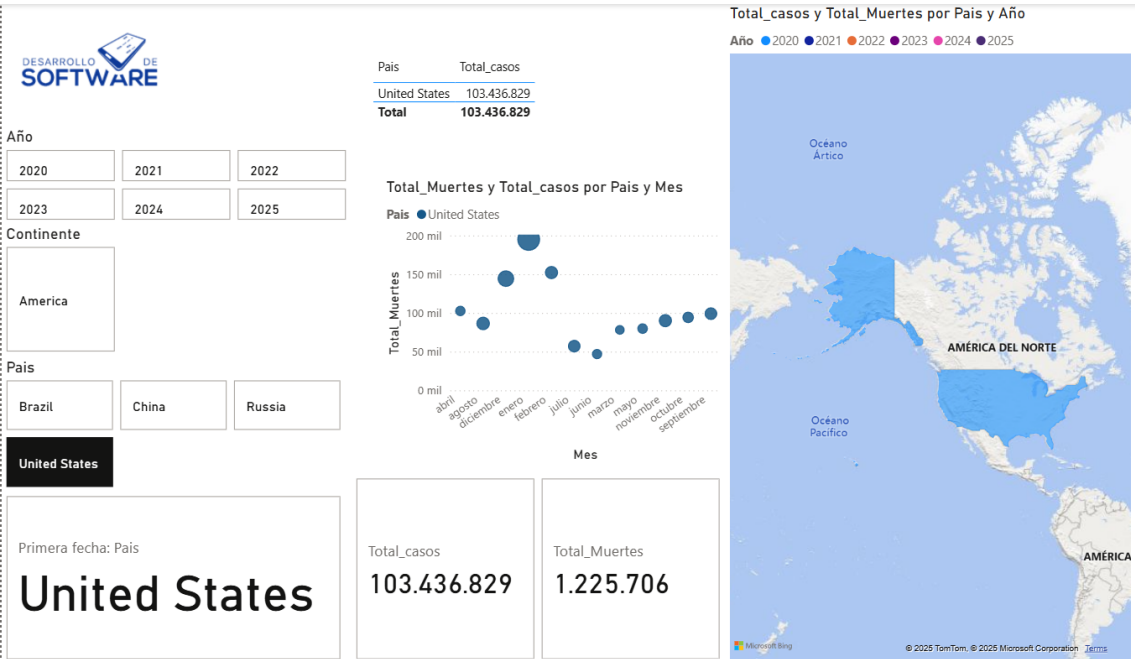
Incluye datos descriptivos de los países, como el nombre del país, su capital y el continente al que pertenece.

Está relacionada con HeCovidMundial por el campo Pais en una relación de uno a muchos.

### **Conclusión del modelo:**

El modelo está diseñado siguiendo la estructura de copo de nieve, separando la información descriptiva (dimensiones) de los datos numéricos (hechos). Este diseño facilita hacer análisis detallados, por ejemplo, ver cómo evolucionaron los casos y las muertes en distintos periodos de tiempo o compararlos entre continentes. Además, al estar normalizado, optimiza el almacenamiento y mejora el rendimiento en consultas y visualizaciones.

5.2. Capturas de pantalla de los resultados obtenidos en los tableros 1 y 2 luego de aplicar el proceso de Minería Manual (KDD) . Establecer conclusiones en cada captura de pantalla.



Comparando los datos de Estados Unidos y Rusia los resultados son los siguientes:

1. Cantidad total de casos
  - Estados Unidos presenta 103.436.829 casos acumulados, cifra que cuadruplica ampliamente los 24.901.467 casos reportados en Rusia.

- Esta diferencia tan marcada puede deberse a la combinación de factores como la población total, las estrategias de testeo, la densidad urbana y la transparencia en el reporte de datos.

## 2. Total de muertes acumuladas

- En términos absolutos, Estados Unidos registra 1.225.706 muertes, mientras que Rusia presenta 404.290 muertes.
- Aunque la cifra de Estados Unidos es mucho mayor, también habría que analizar la tasa de letalidad (muertes/casos) para comparar proporcionalmente.
  - Letalidad aproximada EE.UU.: 1,18%.
  - Letalidad aproximada Rusia: 1,62%.  
Esto sugiere que, proporcionalmente, Rusia tuvo un impacto mayor en relación con la cantidad de infectados.

## 3. Comportamiento mensual

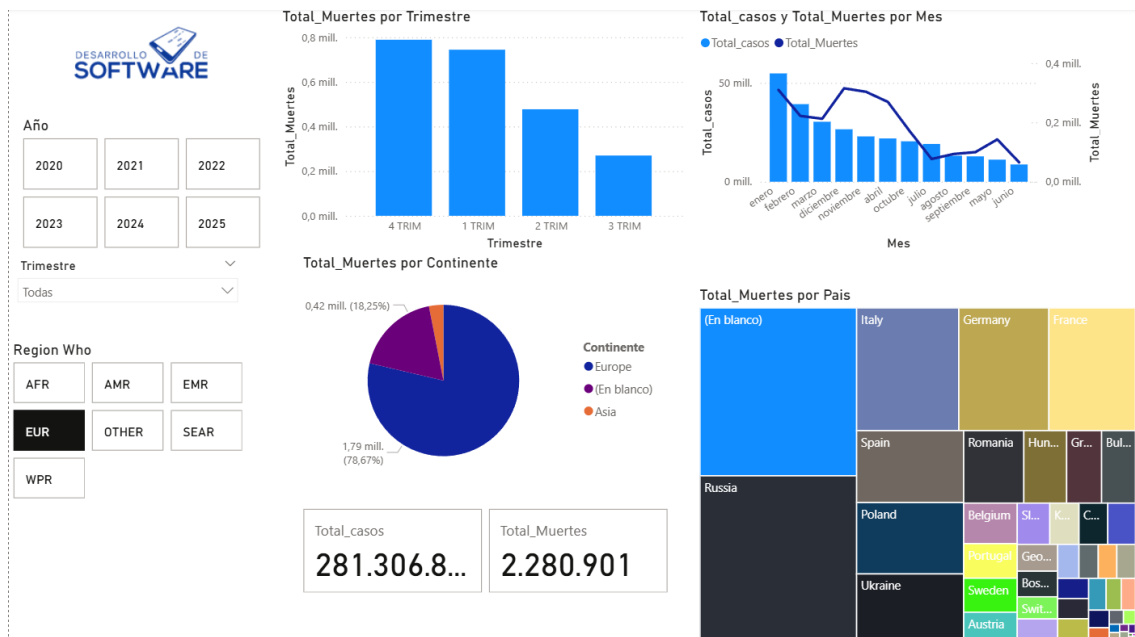
- En Estados Unidos se observan picos claros de muertes en los meses de enero y julio, lo que podría indicar olas específicas de contagio o la aparición de variantes más letales.
- En Rusia, aunque la magnitud de muertes mensuales es menor en valor absoluto, el patrón de variabilidad es más distribuido a lo largo del año, con incrementos notables en febrero, julio y octubre.

## 4. Distribución geográfica y contexto

- El mapa de Estados Unidos destaca su ubicación en América del Norte, mientras que Rusia, al ser el país más extenso del mundo, se visualiza abarcando parte de Europa y Asia.
- Estas diferencias geográficas también influyen en la dispersión de la pandemia y en las estrategias de control.

## 5. Interpretación general

- Estados Unidos fue más afectado en cifras absolutas tanto en casos como en muertes, mientras que Rusia muestra una tasa de letalidad más alta.
- El análisis sugiere que no solo el número de casos es relevante, sino también la capacidad del sistema de salud, las políticas de prevención y la respuesta ante olas de contagio.



### Distribución de muertes por trimestre

- El mayor número de muertes se concentra en el 4º trimestre y el 1º trimestre, ambos con valores cercanos a los 0,7-0,8 millones.
- El 3º trimestre presenta el menor número de muertes, lo que puede estar relacionado con factores estacionales o menor propagación en esos meses.

### Comportamiento mensual de casos y muertes

- Los picos de casos se observan en enero y febrero, coincidiendo con picos altos de muertes.
- A partir de marzo, los casos disminuyen progresivamente, aunque hay un ligero repunte en octubre.

### Muertes por continente

- Europa concentra la gran mayoría de muertes, con 1,79 millones (78,67%) del total registrado.
- Asia y otros continentes tienen un impacto mucho menor en comparación dentro de este filtro.

### Total de casos y muertes

- En total, se registran 281.306.8xx casos y 2.280.901 muertes en la región.
- Esto refleja una tasa de letalidad aproximada del 0,81%, relativamente baja si se compara con otras regiones, pero significativa en valores absolutos.

### Países más afectados

- Dentro de Europa, los países con mayor número de muertes incluyen Italia, Alemania, Francia, España y Rusia.
- Rusia aparece destacada junto a países de Europa Occidental, lo que sugiere que su impacto en muertes es considerable en comparación con su población.