**EVENTO ANIVERSARIO SATI-Q**

**Buceando en los datos SATI-Q sin ahogarse en el intento**

**Experiencia de análisis de datos**

***Coordinadores***: María del Pilar Arias- Ariel L Fernández

*En algún lugar, algo increible está esperando por ser descubierto.* **-Carl Sagan-**

**1-INTRODUCCIÓN**

**1.1 Objetivos**

* Que los participantes adquieran herramientas que les permitan el análisis básico de los datos registrados en el software SATI-Q en su Unidad, de forma que dispongan de información util para la gestión
* Mostrar a los participantes herramientas básicas de visualización de datos que faciliten la comunicación de los resultados en su institución asi como en conferencias científicas.
  1. **¿Qué vamos a aprender?**

***Parte 1***

1. Extraer los datos recolectados con el software SATI-Q usando el sistema de consultas.
2. Transformar los datos.

**Parte 2**

1. Obtener información de los datos extraídos.
2. Realizar visualizaciones y análisis estadísticos básicos.

**2- DESARROLLO DEL TALLER**

**2.1 Extracción de los datos para análisis**

Utilizaremos el **sistema de consultas QCIC** yconstruiremos 4 bases

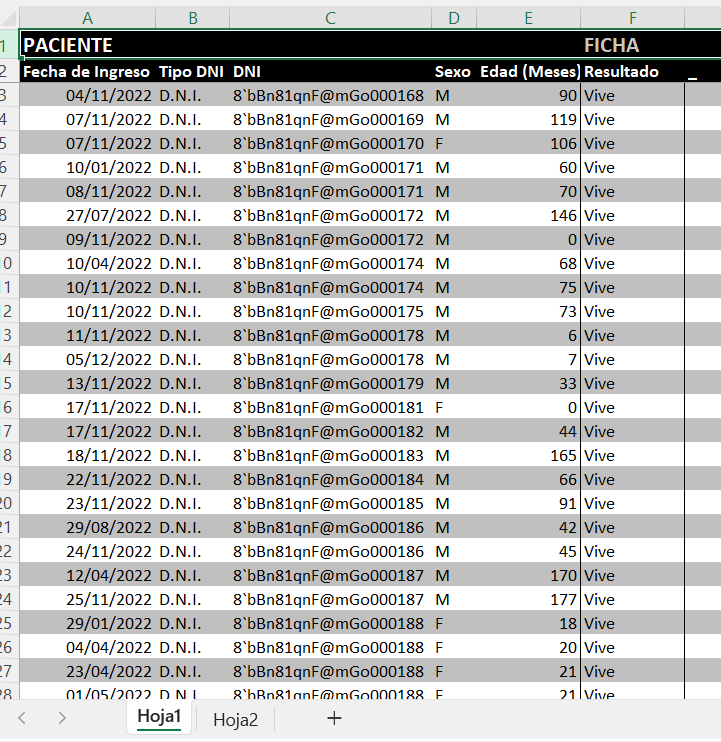
***¡Manos a la obra!***

Exportar los datos por **fecha de ingreso** entre 01-01-2013 y 31-12-2022; seleccionar Tipo paciente Pediátrico, Listar Edad y Resultado

Luego de exportar los datos se genera un archivo en MSExcel que se llama Libro1 por default. Tenemos que eliminar la primera fila, PACIENTE y FICHA para construir la base de análisis (Figura1). **Fecha de ingreso, Tipo de DNI y DNI son las claves de las distintas tablas de SATI-Q.**

El sexo se exporta por default.

***Figura 1***



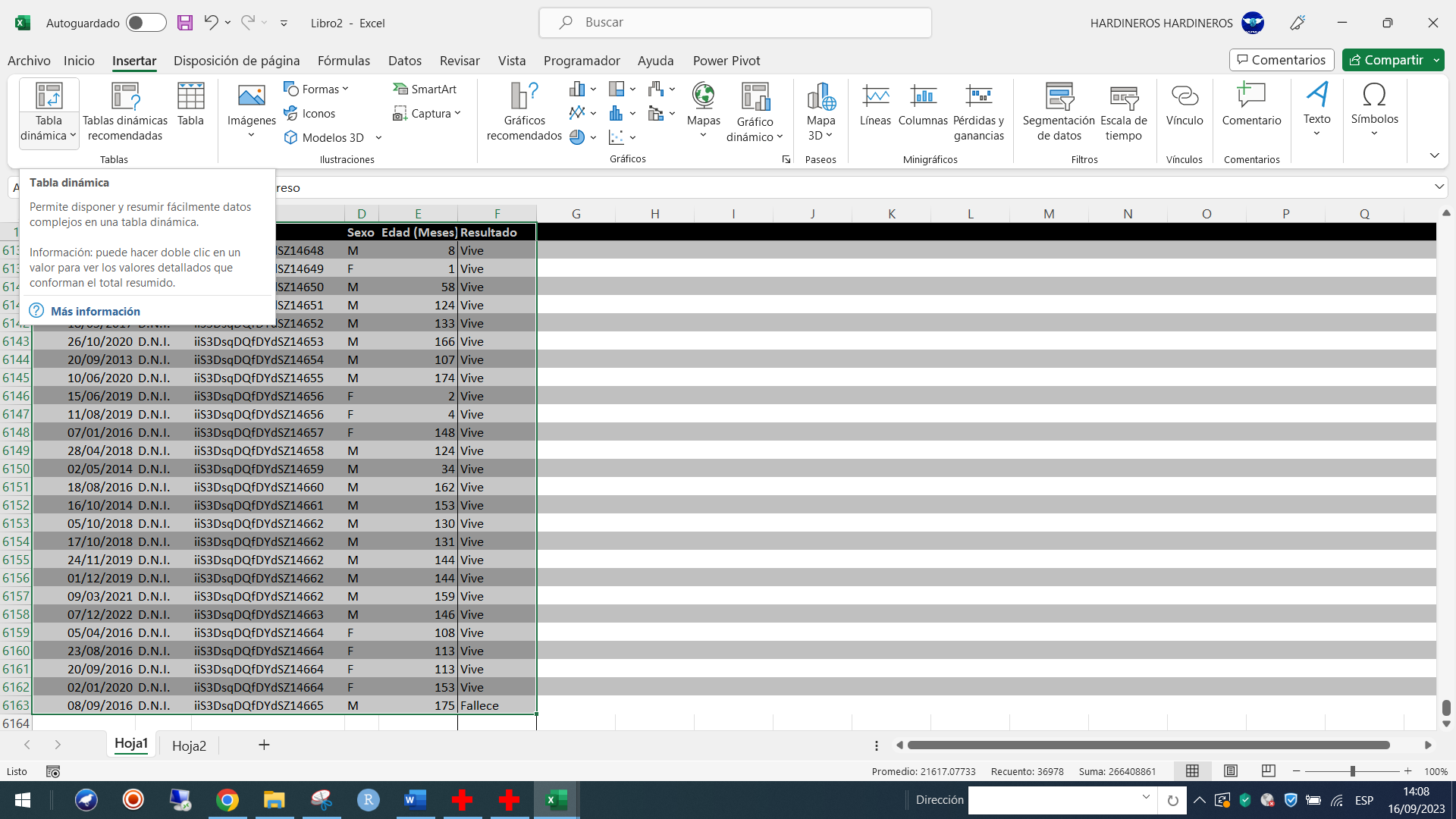
Con estos datos se constituye la base Exportación\_1 con la que vamos a construir una tabla dinámica

***Base Exportación\_1***

1. Tabla dinámica.

Se seleccionan los datos y se indica Insertar Tabla Dinámica (Figura 2) en una nueva hoja

***Figura 2***

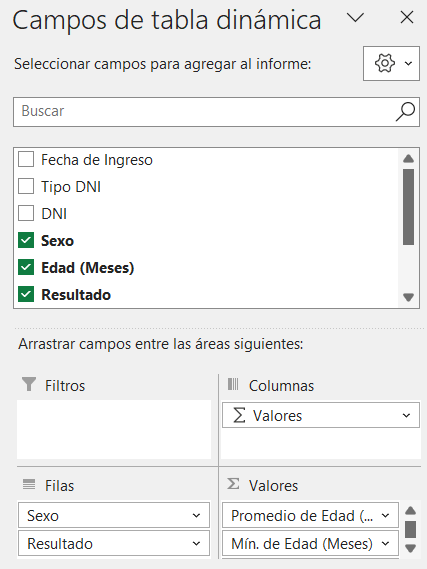


Seleccionando en la Tabla Dinámica los campos Sexo y Edad, veremos cómo se comporta la edad de los pacientes en relación al Sexo.

Calcularemos el promedio de edad, minimo, máximo y desvío estandar.

Luego agregamos el resultado y podemos ver la edad de los que fallecen y la de aquellos que sobreviven, discriminado por Sexo, (Figuras 3 y Tabla 1)

***Figura 3***

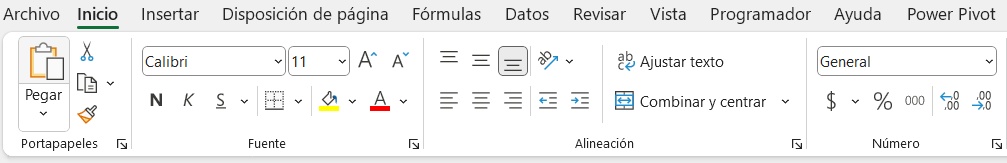


***Tabla 1***



\*\* Opcional: podemos renombrar las columnas y redondear los decimales para mejorar la presentación. (Figura 4, Tabla 2)

**Figura 4**



Modificar decimales

***Tabla 2***

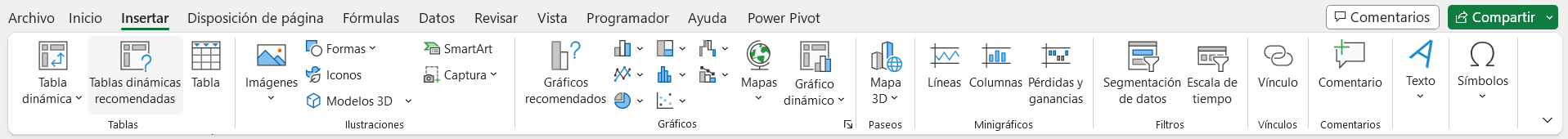


***De acuerdo a la tabla 2, parece ser que los pacientes que fallecen tienen mayor edad que los que sobreviven. ¿Por qué no lo mostramos en un gráfico?***

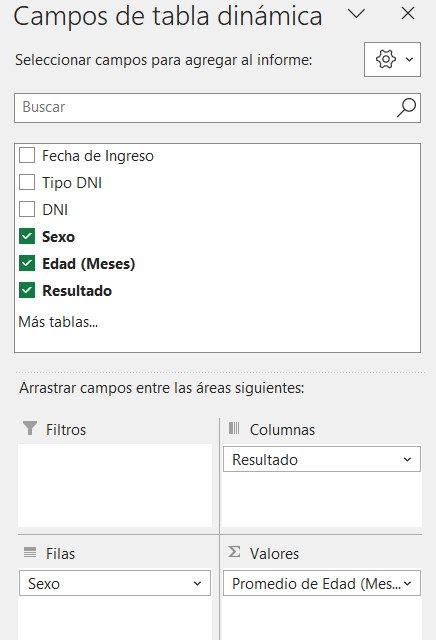
1. Gráfico dinámico

Seleccionamos los datos y luego seleccionamos Insertar gráfico dinámico (Figura 5-7)

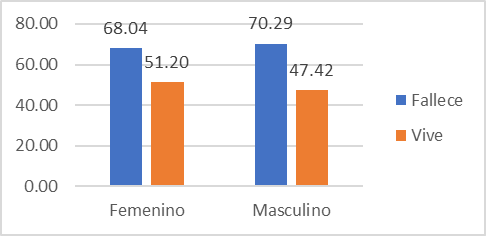
**Figura 5**



***Figura 6***

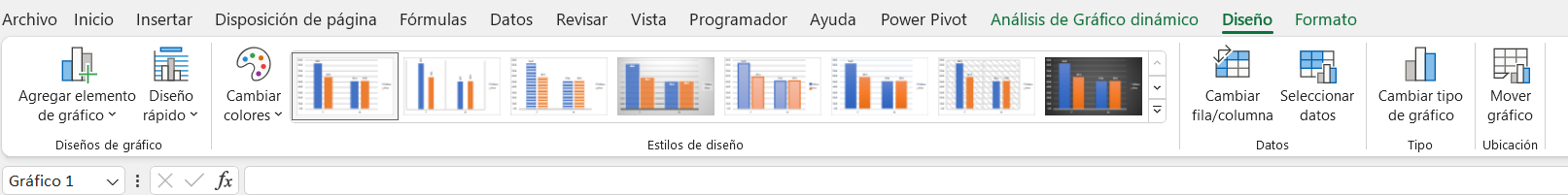


***Figura 7***



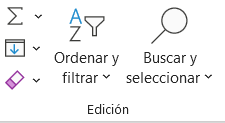
Seleccionando el gráfico y luego la sección Diseño se puede modificar el tipo de gráfico agregar titulos, entre otras funcionalidades (Figura 8)

**Figura 8**



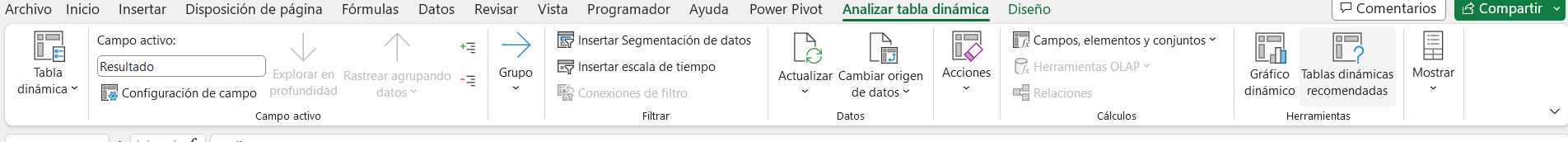
1. Con estos mismos datos podemos buscar otra información como ver la **evolución de los ingresos** en el tiempo. Para eso tengo las fechas de ingreso.

Primero vamos a ORDENAR por columna usando la funcionalidad Ordenar y Filtrar y vemos que los primeros datos corresponden al año 2013 y los últimos al año 2022



Luego vamos a generar las variables Año y la variable Mes usando la **FUNCIÓN AÑO Y MES** de excel. Para incluir estos datos en los análisis necesito **ACTUALIZAR** los datos y CAMBIAR ORIGEN en la tabla o gráfico dinámico (Figura 9)

**Figura 9**

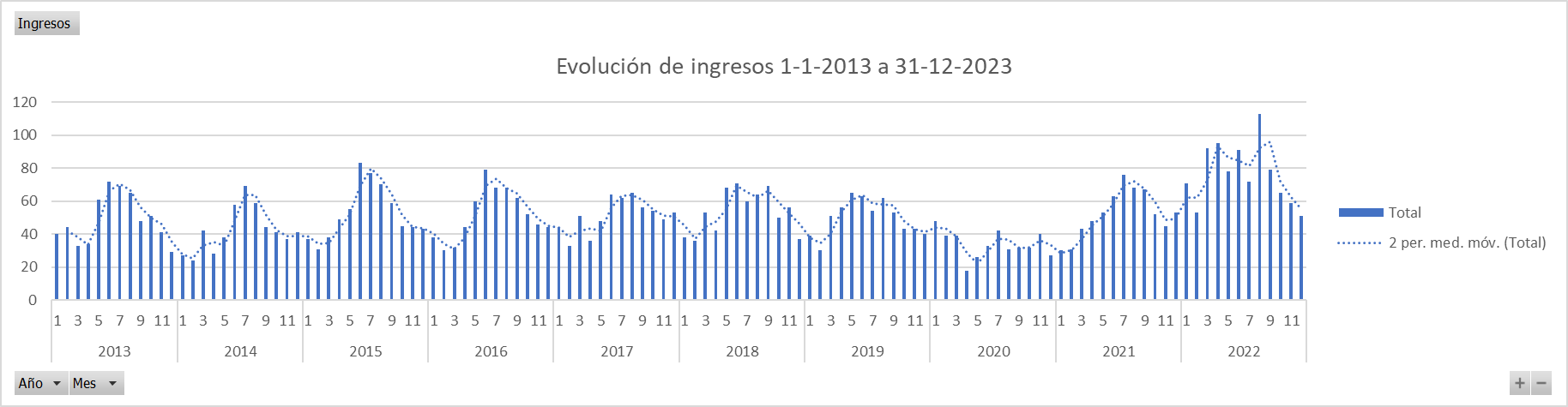


Ahora vamos a hacer un gráfico dinámico incluyendo año y mes, agregaremos título al cuadro, agregaremos linea de tendencia usando media movil y cómo seleccionar el tipo de gráfico (usaremos las funcionalidades DISEÑO DE GRAFICOS RAPIDO/ AGREGAR ELEMENTO AL GRAFICO; ver figura 8). En el gráfico podemos CONTRAER y EXPANDIR (iconos +/-) los datos que se visualizan (Figuras 10-11)

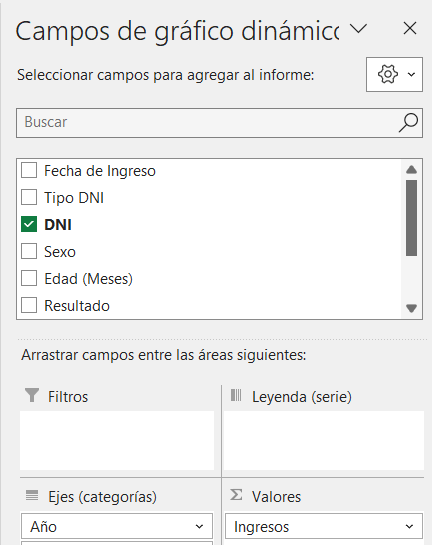
Por último, vamos a GUARDAR el gráfico como IMAGEN (Hoja Imagen)

*Se nota la estacionalidad de los ingresos y el descenso en el año 2020*

**Figura 10**



**Figura 11**



***¡¡Cuanta información obtuvimos mirando muy pocas variables!!***

***Base Exportación\_2***

Usando el Sistema de Consultas se incluirán a los datos previos, los datos de ingreso. Se agregan las variables **Procedencia, Motivo de ingreso, Fecha de ingreso al hospital, Dias de estadía, PIM 3 score y Probabilidad, PIM2 Score y Probabilidad.**

Exportar datos

*Esta base de datos también se puede utilizar para analizar en Stata, R, Phyton entre otros*

Se elimina la fila 1 ya que no es el nombre de las variables, sino que se refiere a las solapas de donde proviene la información.

Vamos a **poner filtros y renombrar las variables**, por ejemplo Fech\_ingreso; Estadía\_UCI; Motivo\_ingreso; Score\_PIM II; Prob\_PIM2; Score\_PIM 3;Prob\_PIM3

***¡Tengo problemas! la probabilidad de muerte cambio por periodo, primero usaba PIM2 y luego PIM3.* Necesito generar una probabilidad de muerte unificada**

GENERAR Columna Probabilidad Prob\_mort usando la función de excel SI

=SI (M2=0; K2;M2)

GENERO Año y Mes

Vemos que la cantidad de datos faltantes disminuyó en la variable probabilidad de muerte

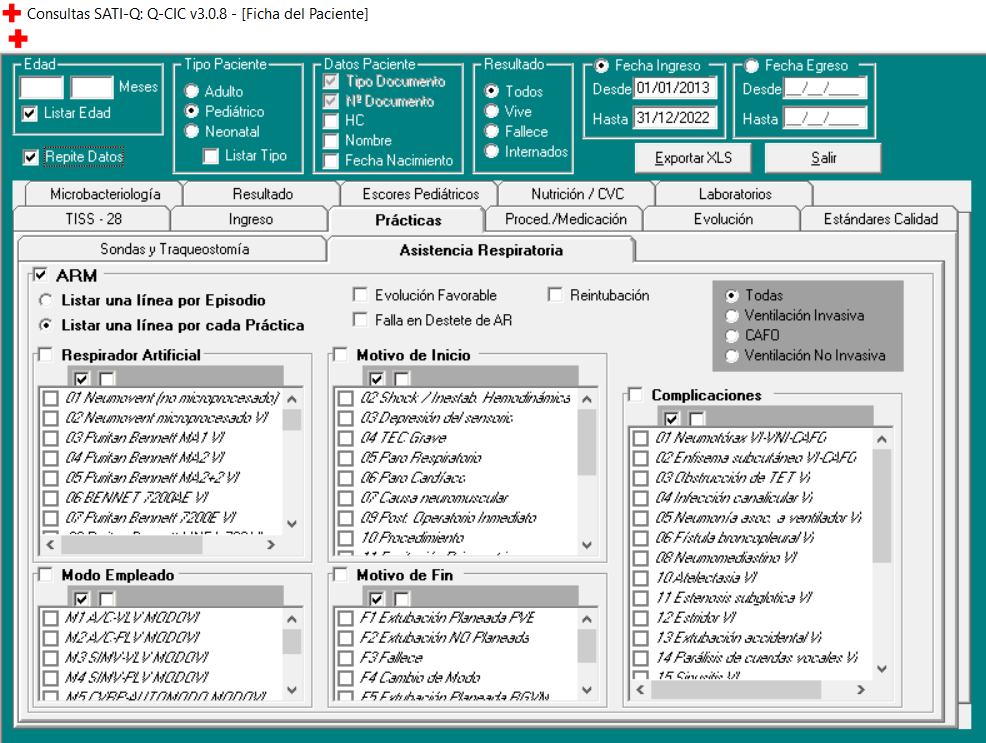
***Base Exportación\_ARM***

Usando el sistema de consultas, exporto los datos correspondientes a la práctica ARM (Figura 12)

**Seleccionar UNA LINEA POR CADA PRACTICA**.

Seleccionar REPITE DATOS y TODAS.

**Figura 12**



Esta exportación me muestra los días de ARM y tipo de apoyo ventilatorio. Las variables son

* Asistencia Respiratoria SI/NO
* Soporte Ventilatorio: ARM/VNI/CAFO. Se muestra cada episodio en una línea distinta
* Fecha Inicio Vent
* Fecha Finalización Vent
* Total Días VI
* Total Días CAFO
* Total Días VNI

*De esta forma tengo más filas que internaciones porque cada vez que un paciente tuvo un tipo distinto de ARM genera una fila*

***Base Exportación\_ARM\_episodio***

**SELECCIONAR Listar una línea por episodio**. (Figura 12). De esta forma el Soporte Ventilatorio se muestra en una sola columna separada por comas

Las variables que tenemos son

* Fecha de Ingreso
* Tipo DNI
* DNI
* Sexo
* Edad (Meses)
* Resultado
* Asistencia Respiratoria SI/NO
* Cantidad AR: cantidad de episodios de Soporte ventilatorio que tuvo el paciente en la internación. Ej si tuvo ARM, VNI y CAFO , el valor de Cantidad de AR será 3
* Soporte Ventilatorio: ARM/VNI/CAFO . Se muestra en una sola linea, separado por coma
* Fecha Inicio Vent
* Fecha Finalización Vent
* Dias\_VI : es la suma los días de VI que tuvo el paciente en la internación
* Días\_CAFO: es la suma de los días de CAFO que tuvo el paciente en la internación
* Días\_VNI: es la suma de los días de VNI que tuvo el paciente en la internación

**Ya tenemos las 4 bases de datos que vamos a analizar** usando el lenguaje de programación Phyton, en el ambiente de Google Colab

**Para la parte 2 vamos a usar la notebook FINAL\_TallerSATI-Q.ipynb**

La plataforma **Google Colab** , permite ejecutar el lenguaje de programación Phyton sin necesidad de una configuración particular del equipo que estamos usando (<https://colab.research.google.com/?hl=es#scrollTo=-Rh3-Vt9Nev9>)

En este enlace pueden ver las características principales de una notebook o cuaderno de Colab : <https://colab.research.google.com/notebooks/basic_features_overview.ipynb>

**Sigamos el práctico en la notebook**