**Trabajo Final de Análisis de Datos**

**Nombre del estudiante: Yeinan del Carmen Uparela Arango**  
**Curso: MUJERES TIC GRUPO PROFE YOHAN RODRIGUEZ**  
**Fecha: 20/052025**

**1. Título del Proyecto**

Análisis de los registros sanitarios de dispositivos médicos y otras tecnologías.  
*Ejemplo*: “Análisis de la satisfacción mundial y su relación con el PIB per cápita”.

**2. Descripción del Dataset**

**Nombre del dataset:** Registros sanitarios de dispositivos médicos y otras tecnologías.

**Fuente del dataset (URL):** <https://www.datos.gov.co/Salud-y-Protecci-n-Social/REGISTROS-SANITARIOS-DE-DISPOSITIVOS-M-DICOS-Y-OTR/y4qt-w6tk/about_data>

**Número de registros y columnas:** 407119 Registros y 20 columnas  
**Descripción general del contenido:**

Esta base de datos contiene los datos actualizados a 16 mayo de 2025 de todos los registros sanitario ante el invima de los dispositivos Medicos que han sido fabricados en Colombia y los que se han importado. En esta hay registros de sus usos, la vida util, el nivel de riesgo, los fabricantes e importadores de estos productos.

Este conjunto de datos nos permite analizar cuáles son los productos mas importados, que cantidades son las que se fabrican en Colombia.

**3. Objetivo del Análisis**

Plantea uno o varios objetivos claros que guíen tu análisis. Este apartado debe responder:

* Determinar la cantidad de productos que se fabrican en Colombia
* Identificar cuales son los productos con más importaciones
* Identificar la ciudad en Colombia que más fabrica estos productos
* Conocer cuales son los productos que más se utilizan
* Observar que tanto fabrica Colombia estos productos, la capacidad de producción.

Mi motivo a explorar esta base de datos es Conocer el mercado y la industria Colombiana de estos productos médicos, que ciudades son las principales fabricantes en Colombia.

Ver la influencia del mercado internacional en Colombia y que tanto son las importaciones de estos productos.

**4. Proceso de Limpieza de Datos**

Describe los pasos que realizaste para preparar el conjunto de datos, tales como:

* Filtrar todos los registros cuya columna “estado de registro” tenía valor “vencido”.
* Eliminar todos los registros de la columna “vida útil” aparezcan con valor NaN o null.
* Eliminar todos los registros de la columna “vida útil” que no especifique si son años o meses.
* Se eliminaron las columnas “observaciones” y “modalidad”.

**5. Transformación de Datos**

Indica qué cambios hiciste para transformar los datos, tales como:

* Todos los registros de la columna “vida útil” fueron transformados a mayúsculas para unificar valores (“años”, “AÑOS”).

Estos datos fueron modificados para un mejor manejo de los datos y asegurarnos que este todos los datos de la misma forma para no omitir ninguna información que afecte el resultado final.

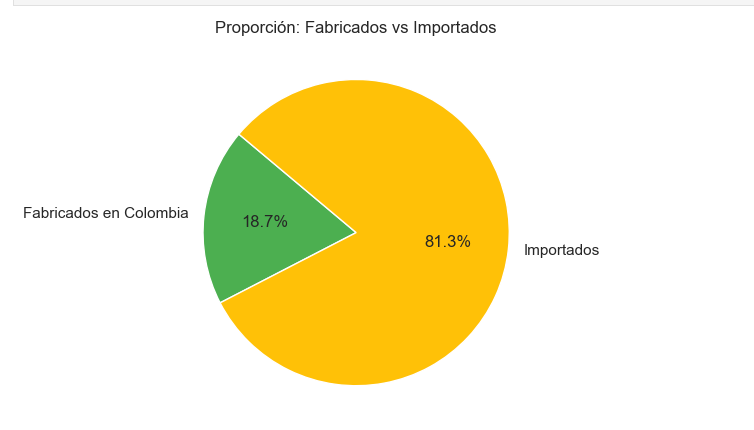
**6. Proceso ETL (Extract, Transform, Load)**

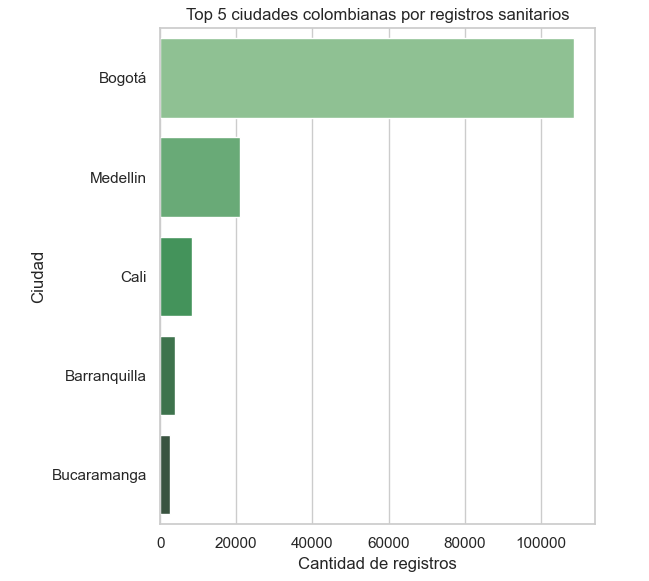
* **Extract:** Esta base de datos fue obtenida de la página de datos abiertos del gobierno de Colombia. Fue una descarga directa de la pagina obtenida en un archivo en formato .csv.
* **Transform:** la base de datos se modifico eliminando los datos vencidos, datos que no estaban definidos y se modifico los datos de la columna “vida útil” de minúscula a mayúscula para facilitar la operación de los datos.
* **Load:** el nuevo Dataset limpio fue almacenado de forma local en Archivo .csv.

**7. Análisis Exploratorio de Datos (EDA)**

Expón el análisis que realizaste utilizando estadísticas descriptivas y visualizaciones. Puedes incluir:

* Distribución de registros por país: Colombia representa aproximadamente el 18% delos productos fabricados del total de registros.



* Ciudades como Bogotá, Medellín, Cali, Barranquilla y Bucaramanga son los principales centros de fabricación.

Se observan correlaciones como:

* Mayor vida útil esta relacionada con el nivel de riesgo.
* Las ciudades con gran numero de registros son las que más tienen empresas fabricantes.

Se utilizo pandas ,matplotlib y Seaborn.

**8. Conclusiones**

Reflexiona sobre los resultados obtenidos:

Puede descubrir que Colombia realiza importaciones de productos tecnológicos para diagnósticos y Bogotá es el principal productor en Colombia, que Colombia fabrica un buen porcentaje de estos productos médicos.

Puede cumplir con los objetivos planteados ya que se identificaron los productos con mayor importación y las ciudades que mas fabrican estos dispositivos a nivel nacional.

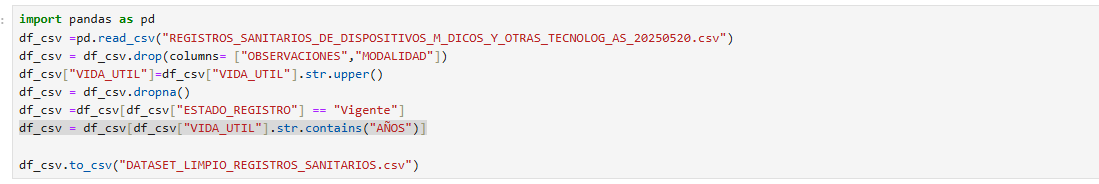
Tuve varias limitaciones a la hora de realizar este informe principalmente al momento de buscar la base de datos con las especificaciones de la cantidad de registros.

Recomendaciones que el Invima registre la cantidad de productos y sus valores para tener mayor claridad del impacto económico que este mercado genera en el país.

**9. Anexos (opcional)**

Incluye enlaces al repositorio de GitHub (si lo creaste), capturas de visualizaciones o código relevante que no hayas incluido en las secciones anteriores.

Con este código genero el dataset limpio.



Código que genera las graficas pie chart donde muestra el porcentaje de fabricación.



Código que genera el diagrama de barra donde se muestra el top 5 de ciudades fabricantes.

