

|  |
| --- |
| Evaluación Sumativa 4 |
| Modulo 3: Transformación y limpieza de datos en Power BI Desktop |
| 9 octubre  Gregory Cuevas Navarrete |

# CONTEXTO DATASET

|  |
| --- |
| Ventas de Articulos Electronicos **Los datos contienen miles de compras en tiendas de electrónica desglosadas por el mes de Abril, tipo de producto, costo, dirección de compra, etc.** Enlace GibHub |
| https://github.com/Gregory-Caves/Evaluaci-n-Sumativa-4.git |
| Capturas de Pantalla  1. Descargar el Dataset:    * Buscar una data con necesidad de limpieza y transformaciones. Descarga el archivo en tu computadora.      1. Abrir Power BI Desktop:    * Abre Power BI Desktop e importa el archivo de módulos anteriores.      1. Panel de Consultas: Vista Previa de los Datos:    * Accede al panel de consultas y revisa la vista previa de los datos. Verifica que los datos se hayan importado correctamente.     Se comienza a revisar previamente la información y desde ya se identifican celdas en blanco y columnas repetidas   1. Eliminar Filas Duplicadas:    * Identifica y elimina cualquier fila duplicada en el dataset.    * Pasos: Selecciona las columnas relevantes para la eliminación de duplicados y usa la opción "Eliminar duplicados".     En este apartado se eliminaron todas aquellas filas que estaban duplicadas y no entregaban un valor importante a la base de datos, así como también se eliminaron aquellas filas en blanco   1. Cambiar Tipos de Datos:    * Cambia los tipos de datos de las columnas para que coincidan con la naturaleza de los datos. Por ejemplo, convierte columnas de texto a formato de fecha o numérico.    * Pasos: Ajusta los tipos de datos desde el menú "Tipo de datos".   Interfaz de usuario gráfica, Aplicación  Descripción generada automáticamente  Se modificó el tipo de la columna order\_id, de numero a tipo texto   1. Quitar Columnas Innecesarias:    * Elimina las columnas que no sean relevantes para el análisis. Por ejemplo, elimina columnas de identificadores internos que no aportan valor al análisis.    * Pasos: Selecciona las columnas innecesarias y usa la opción "Eliminar columnas".     Se eliminaron las columnas received\_by & la columna retired\_by, ya que ambas columnas entregan un valor innecesarios y repetidos.   1. Rellenar Valores Faltantes:    * Rellena los valores faltantes en el dataset utilizando métodos apropiados, como reemplazar con un valor promedio o un valor predeterminado.    * Pasos: Usa las opciones "Rellenar hacia abajo" o "Rellenar hacia arriba", o reemplaza valores nulos manualmente.       En las celdas en blanco se reemplazan por la palabra “SIN DATO” , con el fin de mantener una tabla con información completa   1. Crear Nuevas Columnas o Tablas:    * Crea al menos una nueva columna calculada o una nueva tabla basada en los datos existentes. Por ejemplo, crea una columna que calcule el total de ventas incluyendo impuestos.    * Pasos: Usa la opción "Agregar columna" y aplica fórmulas para calcular nuevos valores.       Finalmente se crea una nueva columna, teniendo el resultado de Venta Total, multiplicando item\_price \* quantity. |