

# Documentación proyecto online retail sales / Project Documentation Online Retail Sales

Este proyecto consiste en un análisis completo de datos de principio a fin utilizando Power BI, descrito en los siguientes pasos:

***This project consists of a comprehensive end-to-end data analysis using Power BI, outlined in the following steps:***

- 1. Carga de Datos:** Cargamos nuestra base de datos en Power BI, que en este caso era un archivo CSV.  
***Data Loading: We loaded our dataset into Power BI, which in this case was a CSV file.***
- 2. Abrir el Editor de Power Query:** Accedimos al Editor de Power Query para comenzar con la limpieza de datos.  
***Opening Power Query Editor: We accessed Power Query Editor to start cleaning the data.***
- 3. Eliminación de Filas Duplicadas:** Eliminamos cualquier fila duplicada, aunque en este caso no había duplicados presentes.  
***Removing Duplicates: We eliminated any duplicate rows, although no duplicates were present in this case.***
- 4. Revisión de Valores Nulos y Caracteres Especiales:** Verificamos si existían valores nulos, en blanco o con caracteres especiales.  
***Null and Special Character Check: We verified if there were any null, blank, or special character values.***
- 5. Validación de Tipo de Datos y Rango de Fechas:** Revisamos que los tipos de datos fueran correctos según la información de las columnas. También verificamos el rango de fechas, que iba desde el 01/01/2023 a las 12:00 a.m. hasta el 01/01/2023 a las 4:39 p.m., lo que confirmó que los datos representaban un día de ventas.  
***Data Type and Date Range Validation: We ensured the data types were accurate according to the column information. We also checked the date range, which started on 01/01/2023 at 12:00 a.m. and ended on 01/01/2023 at 4:39 p.m., confirming the data represented a single day of sales.***
- 6. Formato de la Columna de Precio:** Cambiamos el tipo de dato de la columna 'Price' de número a moneda y eliminamos los decimales innecesarios.  
***Price Column Formatting: We changed the data type of the 'Price' column from number to currency and removed unnecessary decimals.***
- 7. Eliminación de la Columna de Descuento:** Se eliminó la columna 'Discount', ya que no era relevante para este análisis en particular.

***Removing the Discount Column: The 'Discount' column was deleted, as it was not relevant for this specific analysis.***

8. **Análisis de la Columna de Género:** Revisamos la columna 'Gender' y encontramos que la mayoría de los valores eran "Other" (33%), seguidos de "Female" (33%) y "Male" (32%). "Other" incluye a clientes que no se identifican con géneros tradicionales (hombre o mujer), como personas no binarias o de género fluido.
- Gender Column Analysis: We reviewed the 'Gender' column and found that the majority of values were "Other" (33%), followed by "Female" (33%) and "Male" (32%). "Other" includes customers who do not identify with traditional genders (male or female), such as non-binary or gender-fluid individuals.***

9. **Agrupación de Rangos de Edad:** Para crear visualizaciones más claras, agrupamos las edades de los clientes en una nueva columna con las siguientes categorías:
- Joven Adulto: 19 a 35 años
  - Edad Media: 36 a 52 años
  - Mayor: 53 a 69 años

***Grouping Age Ranges: To create clearer visualizations, we grouped customer ages into a new column with the following categories:***

***Young Adult: 19 to 35 years***

***Middle-Aged: 36 to 52 years***

***Senior: 53 to 69 years***

10. **Visualizaciones:** Una vez que la base de datos estaba limpia y organizada, procedimos a crear las visualizaciones.

***Visualizations: After cleaning and organizing the dataset, we proceeded to create visualizations.***

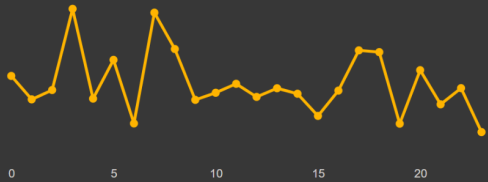
ONLINE RETAIL SALES  
DASHBOARD

\$947 mill.  
Total Sales

Category  
Todas

Continent  
Todas

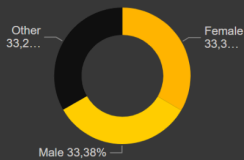
Hours



Category



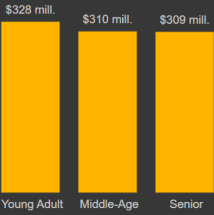
Gender



Location



Age



Payment Method

