

Laboratorio 15

Sesión # 15 Limpieza con Power BI

Título del Laboratorio: Aplicación y uso de la herramienta de Power BI para hacer la limpieza de los datos.

Duración: 2 horas

Objetivos del Laboratorio:

1. *Afianzar los conocimientos y manejo básico en Power BI para realizar las visualizaciones con ejercicios prácticos planteados.*

Materiales Necesarios:

1. *Computador con acceso a internet.*
2. *Colocarlo en el repositorio de Github*
3. *Ampliar el conocimiento con el curso de datos en AWS y Cisco.*
4. *Power BI descargarlo*

Estructura del Laboratorio:

Parte 1

En la primera parte se aplicarán los temas vistos en la sesión como es la limpieza de los datos, se deberá realizar el paso a paso con las respectivas capturas de pantalla, esta aplicación es de acuerdo con los escenarios planteados.

1. Ejercicio de práctica 1.

1. Escenario datos de ventas en línea

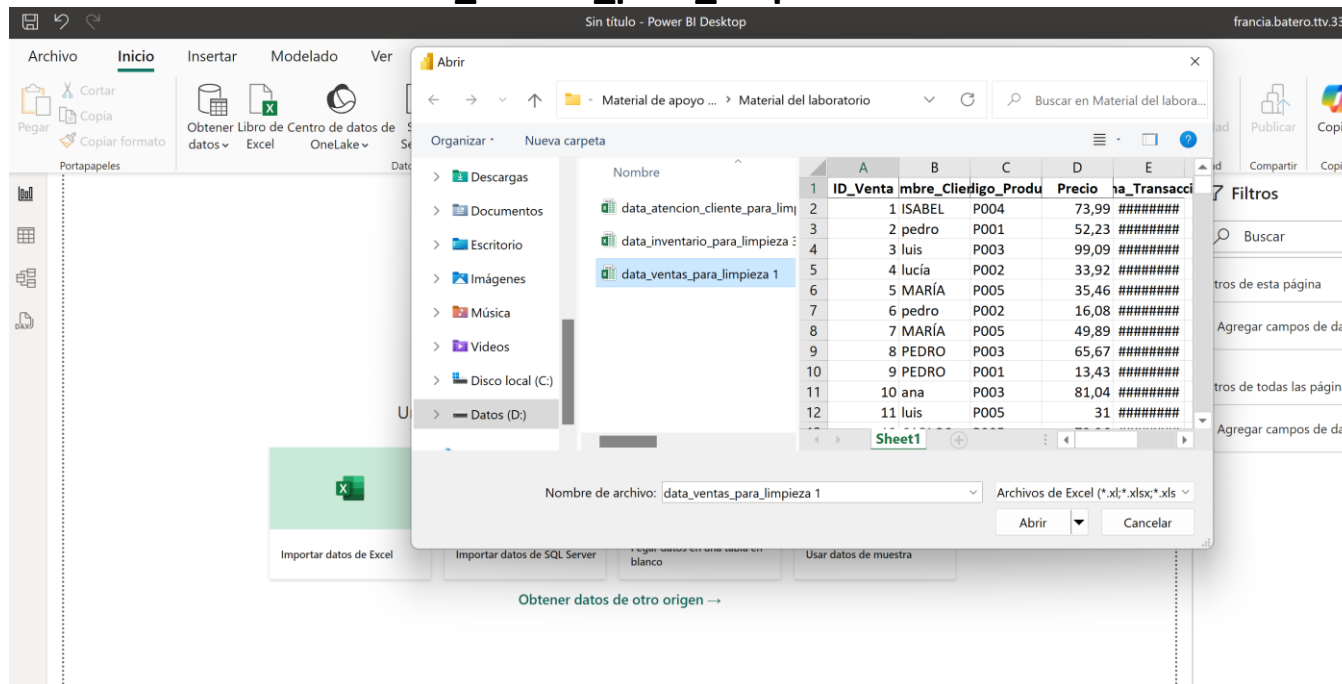
Se tiene un conjunto de datos de una empresa de ventas online que contiene información de clientes, productos vendidos y fechas de transacciones. El archivo tiene problemas comunes como valores faltantes en la columna de códigos de productos, precios con formato incorrecto y nombres de clientes en mayúsculas y minúsculas mezcladas.

Data: data_ventas_para_limpieza 1

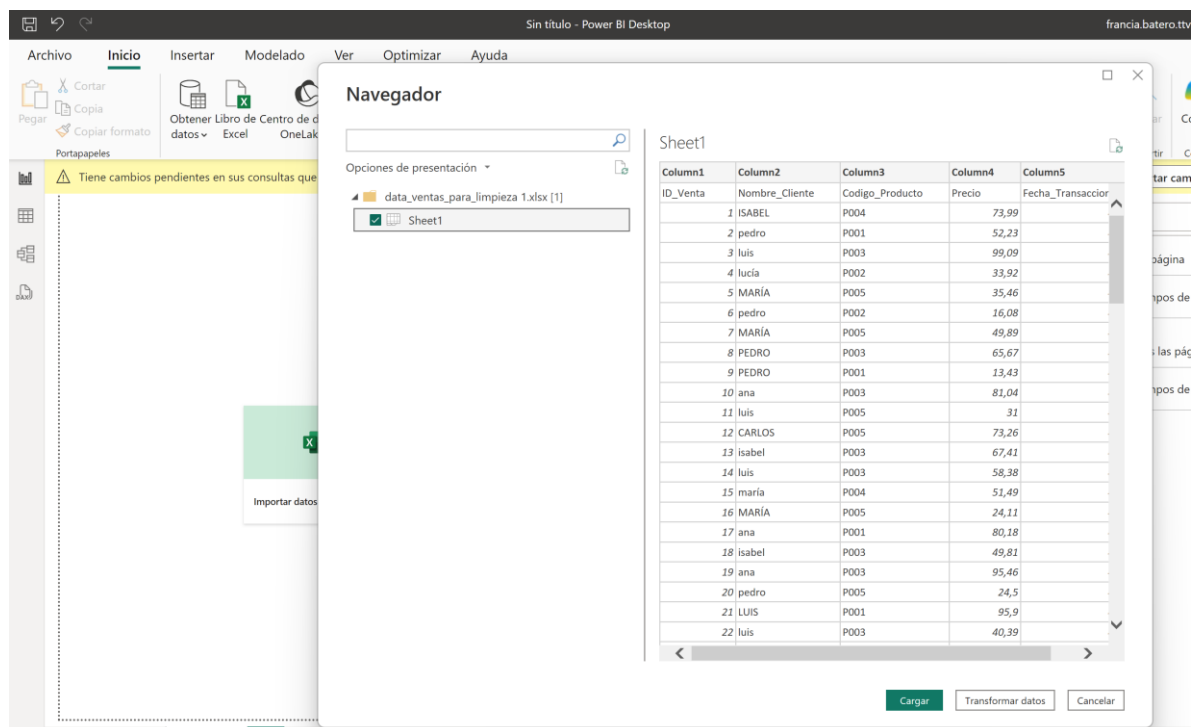
Paso a paso para la limpieza

CONCLUSIONES

- Conectar los datos data_ventas_para_limpieza 1



- Transformamos los datos



Sin título - Editor de Power Query

Archivo Inicio Transformar Agregar columna Vista Herramientas Ayuda

Configuración de la consulta

Barra de fórmulas

Monoespaciada

Distribución de columnas

Mostrar espacio en blanco

Perfil de columna

Calidad de columnas

Permitir siempre

Ir a columna

Editor avanzado

Dependencias de la consulta

Dependencias

Consultas [2]

Venta_tienda_01

Sheet1

= Table.TransformColumnTypes("#Encabezados promovidos",{"ID_Venta", Int64.Type}, {"Nombre_Cliente", type text}, {"Codigo_Producto", type

	ID_Venta	Nombre_Cliente	Codigo_Producto	Precio	Fecha_Transaccion
1	1	ISABEL	P004	73,99	44927
2	2	pedro	P001	52,23	44928
3	3	luis	P003	99,09	44929
4	4	lucía	P002	33,92	44930
5	5	MARÍA	P005	35,46	44931
6	6	pedro	P002	16,08	44932
7	7	MARÍA	P005	49,89	44933
8	8	PEDRO	P003	65,67	44934
9	9	PEDRO	P001	13,43	44935
10	10	ana	P003	81,04	44936
11	11	luis	P005	31	44937
12	12	CARLOS	P005	73,26	44938
13	13	isabel	P003	67,41	44939
14	14	luis	P003	58,38	44940
15	15	maría	P004	51,49	44941
16	16	MARÍA	P005	24,11	44942
17	17	ana	P001	80,18	44943
18	18	isabel	P003	49,81	44944
19	19	ana	P003	95,46	44945
20	20	pedro	P005	24,5	44946
21	21	LUIS	P001	95,9	44947
22	22	luis	P003	40,39	44948
23	23	maría	P005	79,22	44949

• Empezamos a limpiar los Datos

- ✓ Revisar los valores, nulos, duplicados, modificar formatos
- ✓ Cambiamos el formato a mayuscula

Archivo Inicio Transformar Agregar columna Vista Herramientas Ayuda

Columna a partir de los ejemplos * Columna personalizada Invocar función personalizada

Columna condicional

Columna de índice

Duplicar columna

General

Formato

Analizar

Combinar columnas

Extraer

Estadísticas

Estándar

Científico

De número

Trigonometría

Redondeo

Información

De fecha y hora

Fecha

Hora

Duración

Text

Visión Azu

Analytics

Conclusiones d

Consultas [2]

Venta_tienda_01

Sheet1

= Table.TransformColumnTypes("#Encabezados promovidos",{"ID_Venta", type text}, {"Nombre_Cliente", type text}, {"

	ID_Venta	Nombre_Cliente	Codigo_Producto	Precio	Fecha_Transaccion
1	1	ISABEL	P004	73,99	44927
2	2	pedro	P001	52,23	44928
3	3	luis	P003	99,09	44929
4	4	lucía	P002	33,92	44930
5	5	MARÍA	P005	35,46	44931
6	6	pedro	P002	16,08	44932
7	7	MARÍA	P005	49,89	44933
8	8	PEDRO	P003	65,67	44934
9	9	PEDRO	P001	13,43	44935
10	10	ana	P003	81,04	44936
11	11	luis	P005	31	44937
12	12	CARLOS	P005	73,26	44938
13	13	isabel	P003	67,41	44939
14	14	luis	P003	58,38	44940
15	15	maría	P004	51,49	44941
16	16	MARÍA	P005	24,11	44942
17	17	ana	P001	80,18	44943
18	18	isabel	P003	49,81	44944
19	19	ana	P003	95,46	44945
20	20	pedro	P005	24,5	44946
21	21	LUIS	P001	95,9	44947
22	22	luis	P003	40,39	44948
23	23	maría	P005	79,22	44949
24	24	ANA	P003	77,06	44950
25	25	lucía	P004	65,39	44951
26	26	maría		null	44952
27	27	maría	P005	82,57	44953
28	28	MARÍA		null	44954

✓ Filtramos o eliminar valores NULL

Sin título - Editor de Power Query

Archivo Inicio Transformar Agregar columna Vista Herramientas Ayuda

Transponer Tipo de datos: Texto Reemplazar los valores Anular dinamización de columnas
 Invertir filas Detectar tipo de datos Rellenar Mover
 Contar filas Cambiar nombre Columna dinámica Convertir en lista

Dividir columna Formato Extraer Analizar

Combinar columnas Estadísticas Es

Consultas [2]

Venta_tienda_01

Sheet1

Table.TransformColumnTypes(#"Columnas quitadas",{"Codigo_Producto", type text})

ID_Venta	Nombre_Cliente	Codigo_Producto	Precio	Fecha_Transaccion
1			73,99	44927
2			52,23	44928
3			99,09	44929
4			33,92	44930
5			35,46	44931
6			16,08	44932
7			49,89	44933
8			65,67	44934
9			13,43	44935
10			81,04	44936
11			31	44937
12			73,26	44938
13			67,41	44939
14			58,38	44940
15			51,49	44941
16			24,11	44942
17			80,18	44943
18	ISABEL	P003	49,81	44944
19	ANA	P003	95,46	44945
20	PEDRO	P005	24,5	44946
21	LUIS	P001	95,9	44947
22	LUIS	P003	40,39	44948
23	MARÍA	P005	79,22	44949
24	ANA	P003	77,06	44950
25	LUCÍA	P004	65,39	44951
26	MARÍA		28,87	44952
27	MARÍA	P005	82,57	44953

Orden ascendente
 Orden descendente
 Borrar orden
 Borrar filtro
 Quitar vacíos
 Filtros de texto

Buscar

(Seleccionar todo)
 (nulo)
 P001
 P002
 P003
 P004
 P005

Aceptar Cancelar

✓ Cambiar el tipo de datos a cada columna

Sin título - Editor de Power Query

Archivo Inicio Transformar Agregar columna Vista Herramientas Ayuda

Transponer Tipo de datos: Fecha Reemplazar los valores Anular dinamización de columnas
 Invertir filas Detectar tipo de datos Rellenar Mover
 Contar filas Cambiar nombre Columna dinámica Convertir en lista

Dividir columna Formato Extraer Analizar

Combinar columnas Estadísticas Estándar Científico Redondeo Trigonometría
 Información Fecha

Consultas [2]

Venta_tienda_01

Sheet1

Table.TransformColumnTypes(#"Filas filtradas",{"Fecha_Transaccion", type date},{"Precio", type number})

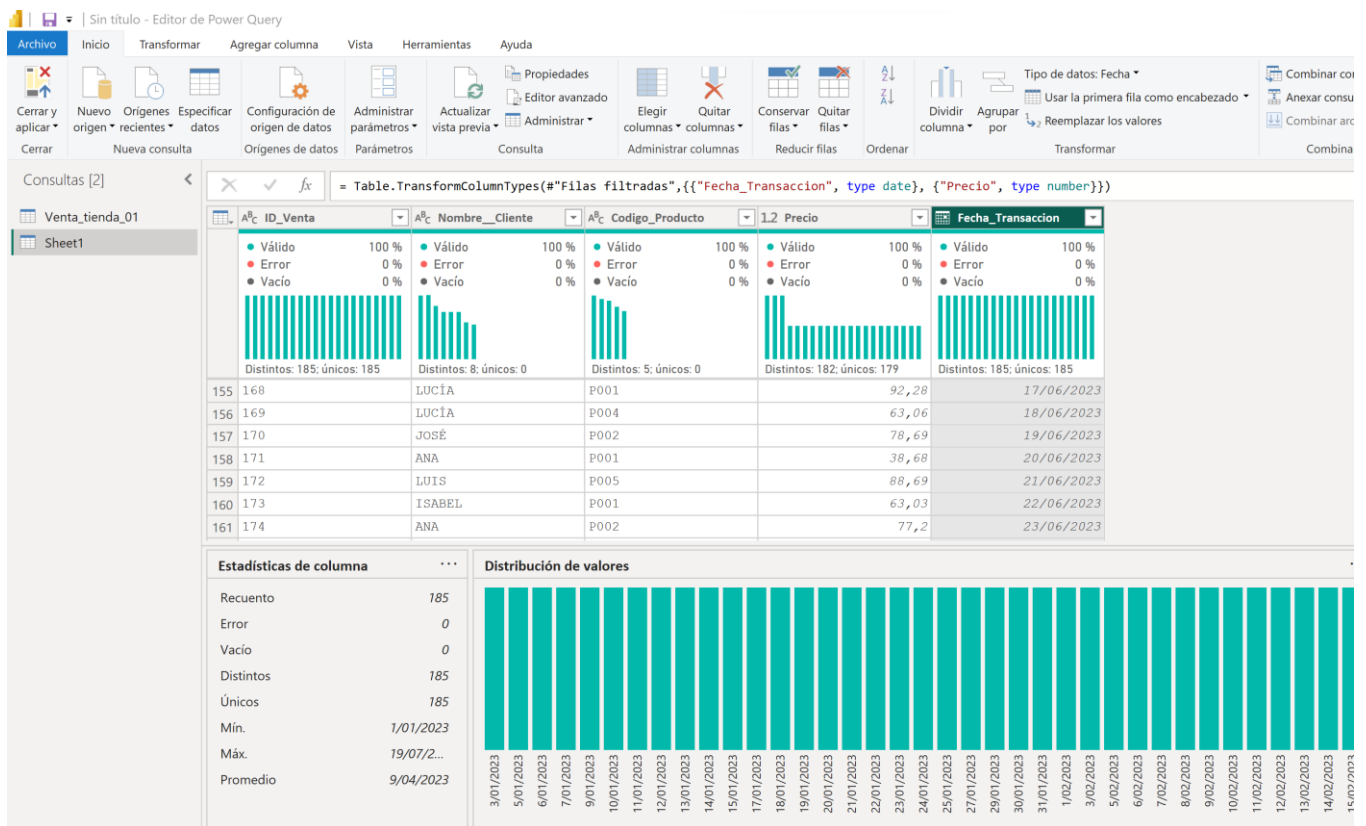
ID_Venta	Nombre_Cliente	Codigo_Producto	Precio	Fecha_Transaccion
155	LUCÍA	P004	92,28	17/06/2023
156	LUCÍA	P004	63,06	18/06/2023
157	JOSÉ	P002	78,69	19/06/2023
158	ANA	P001	38,68	20/06/2023
159	LUIS	P005	88,69	21/06/2023
160	ISABEL	P001	63,03	22/06/2023
161	ANA	P002	77,2	23/06/2023

Estadísticas de columna

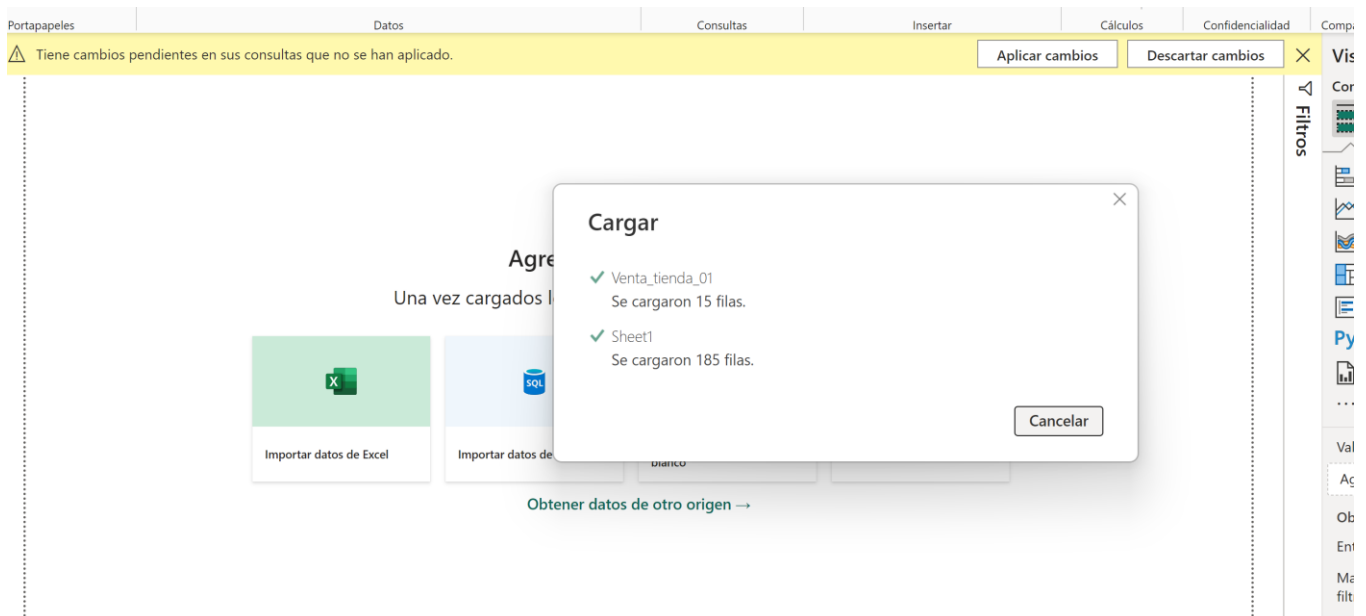
Recuento	185
Error	0
Vacio	0
Distintos	185
Únicos	185
Mín.	1/01/2023
Máx.	19/07/2...
Promedio	9/04/2023

Distribución de valores

- Cerrar y guardar



- Aplicar y cargar Cambios



- Verificación de los datos

ID_Venta	Nombre_Cliente	Código_Producto	Precio	Fecha_Transacción
1	ISABEL	P004	73.99	01/01/2023
2	PEDRO	P001	52.23	02/01/2023
3	LUIS	P003	99.09	03/01/2023
4	LUCÍA	P002	33.92	04/01/2023
5	MARÍA	P005	35.46	05/01/2023
6	PEDRO	P002	16.08	06/01/2023
7	MARÍA	P005	49.89	07/01/2023
8	PEDRO	P003	65.67	08/01/2023
9	PEDRO	P001	13.43	09/01/2023
10	ANA	P003	81.04	10/01/2023
11	LUIS	P005	31.00	11/01/2023
12	CARLOS	P005	73.26	12/01/2023
13	ISABEL	P003	67.41	13/01/2023
14	LUIS	P003	58.38	14/01/2023
15	MARÍA	P004	51.49	15/01/2023
16	MARÍA	P005	24.11	16/01/2023
17	ANA	P001	80.18	17/01/2023
18	ISABEL	P003	49.81	18/01/2023
19	ANA	P003	95.46	19/01/2023
20	PEDRO	P005	24.50	20/01/2023
21	LUIS	P001	95.90	21/01/2023
22	LUIS	P003	40.39	22/01/2023
23	MARÍA	P005	79.22	23/01/2023
24	ANA	P003	77.06	24/01/2023
25	LUCÍA	P004	65.39	25/01/2023
27	MARÍA	P005	82.57	27/01/2023
29	LUCÍA	P002	12.89	29/01/2023

- Conectar los Datos: Importar el archivo "data_ventas_para_limpieza 1" a Power BI.
- Ingresar al Editor de Power BI: Abrir el Editor de Power Query para comenzar las transformaciones.
- Transformar los Datos: Realizar transformaciones iniciales, como revisar nulos y duplicados.
- Limpiar los Datos:
 - Revisar valores nulos y duplicados.
 - Modificar formatos (convertir a mayúsculas).
 - Filtrar o eliminar valores NULL.
 - Cambiar el tipo de datos a cada columna.
- Cerrar y Guardar: Aplicar las transformaciones y regresar a Power BI.
- Aplicar y Cargar Cambios: Confirmar que los datos transformados se carguen correctamente en Power BI.
- Verificación de los Datos: Validar que los cambios sean correctos y que los datos estén listos para el análisis.