

LABORATORIO #15

PARTE 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

ID_Venta	Nombre_Cliente	Codigo_Producto	Precio	Fecha_Transaccion
1	Isabel	P004	73,99	domingo, 1 de enero de 2023
2	Pedro	P001	52,23	lunes, 2 de enero de 2023
3	Luis	P003	99,09	martes, 3 de enero de 2023
4	Lucía	P002	33,92	miércoles, 4 de enero de 2023
5	María	P005	35,46	jueves, 5 de enero de 2023
6	Pedro	P002	16,08	viernes, 6 de enero de 2023
7	María	P005	49,89	sábado, 7 de enero de 2023
8	Pedro	P003	65,67	domingo, 8 de enero de 2023
9	Pedro	P001	13,43	lunes, 9 de enero de 2023
10	Ana	P003	81,04	martes, 10 de enero de 2023
11	Luis	P005	31	miércoles, 11 de enero de 2023
12	Carlos	P005	73,26	jueves, 12 de enero de 2023
13	Isabel	P003	67,41	viernes, 13 de enero de 2023
14	Luis	P003	58,38	sábado, 14 de enero de 2023
15	María	P004	51,49	domingo, 15 de enero de 2023
16	María	P005	24,11	lunes, 16 de enero de 2023
17	Ana	P001	80,18	martes, 17 de enero de 2023
18	Isabel	P003	49,81	miércoles, 18 de enero de 2023
19	Ana	P003	95,46	jueves, 19 de enero de 2023
20	Pedro	P005	24,5	viernes, 20 de enero de 2023
21	Luis	P001	95,9	sábado, 21 de enero de 2023
22	Luis	P003	40,39	domingo, 22 de enero de 2023
23	María	P005	79,22	lunes, 23 de enero de 2023
24	Ana	P003	77,06	martes, 24 de enero de 2023
25	Lucía	P004	65,39	miércoles, 25 de enero de 2023
26	María	Sin codigo	28,87	jueves, 26 de enero de 2023
27	María	P005	82,57	viernes, 27 de enero de 2023
28	María	Sin codigo	70,53	sábado, 28 de enero de 2023
29	Lucía	P002	12,89	domingo, 29 de enero de 2023

CONCLUSIONES: Al abrir el archivo y ver en la tabla datos faltantes o errores en ellos, se hace una limpieza de datos, en el ítem de consultas “Transformar datos” nos permite hacer la debida limpieza, tenemos dos opciones, tanto reemplazar cómo filtrar. Manejamos valores faltantes en código de producto, cambiamos formato en las diferentes columnas de precio, ID de venta y fecha. Finalmente, normalizamos el formato de fuente del nombre del cliente.

2. Escenario: Limpieza de datos de un servicio de atención al cliente Trabajas en un centro de atención al cliente y tienes un conjunto de datos que contiene información sobre las interacciones con los clientes. Este dataset incluye el ID del cliente, Nombre del cliente, Categoría de consulta, Duración de la llamada y Fecha de la interacción. Algunos registros tienen problemas como nombres en minúsculas y mayúsculas inconsistentes, valores faltantes en la categoría de la consulta, y duraciones de llamada con errores de formato.

Data: data_atencion_cliente_para_limpieza 2

	A ^B _C ID_Cliente	A ^B _C Nombre_Cliente	A ^B _C Categoria_Conulta	A ^B _C Duracion_Llamada	Fecha_Interaccion
1	2	José	Soporte Técnico	24.36	2/01/2023
2	3	Carlos	Facturación	58.86	3/01/2023
3	4	José	Facturación	33.57	4/01/2023
4	5	Pedro	Facturación	18.28	5/01/2023
5	7	Pedro	Soporte Técnico	29.0	7/01/2023
6	8	Luis	Consulta General	39.77	8/01/2023
7	9	Lucía	Soporte Técnico	23.21	9/01/2023
8	10	Ana	Consulta General	3.2	10/01/2023
9	11	María	Quejas	2.37	11/01/2023
10	12	José	Soporte Técnico	49.46	12/01/2023
11	13	Isabel	Quejas	26.74	13/01/2023
12	14	Pedro	Soporte Técnico	4.82	14/01/2023
13	15	Isabel	Consulta General	42.04	15/01/2023
14	16	Luis	Quejas	20.38	16/01/2023
15	17	Carlos	Facturación	19.76	17/01/2023
16	18	Isabel	Quejas	51.29	18/01/2023
17	19	Ana	Facturación	10.43	19/01/2023
18	20	María	Sugerencias	31.22	20/01/2023
19	21	Luis	Soporte Técnico	21.74	21/01/2023
20	22	Pedro	Quejas	48.28	22/01/2023
21	23	Isabel	Soporte Técnico	28.45	23/01/2023
22	24	Pedro	Soporte Técnico	23.27	24/01/2023
23	25	Luis	Facturación	39.42	25/01/2023
24	26	Carlos	Consulta General	18.8	26/01/2023
25	27	Luis	Quejas	39.98	27/01/2023
26	28	Lucía	Quejas	16.79	28/01/2023
27	29	Carlos	Sugerencias	32.1	29/01/2023
28	30	Carlos	Sugerencias	17.72	30/01/2023
29	31	Ana	Quejas	23.5	31/01/2023
30	32	Lucía	Soporte Técnico	57.76	1/02/2023

CONCLUSIONES: Al cargar el archivo a power Bi empezamos limpiando, cambiando el formato de las diferentes columnas, eliminando duplicados y reemplazando valores categorizados como errores, tipo “null” por medio del filtrado de filas.

3. Escenario: Tienda minoritaria de tecnología Trabajas para una tienda minoritaria tecnología, el equipo de inventario tiene dificultades para gestionar adecuadamente los datos de productos debido a la calidad inconsistente de los registros. El conjunto de datos de inventario incluye información sobre ID de producto, Nombre del producto, Cantidad disponible, Precio unitario y Fecha de ingreso al inventario. Sin embargo, se han detectado varios problemas en los datos que impiden su uso eficaz para análisis de ventas y reposición de stock. Los problemas para solucionar incluyen:

- Nombres de productos faltantes.
- Precios de algunos productos faltantes.
- Cantidades de productos con texto adicional "unidades".

Data: data_inventario_para_limpieza 3

CONCLUSIONES: En esta limpieza, realizamos varios pasos como, quitar errores denominados “null” filtrar filas erróneas, cambiar el formato de ID, fecha y cantidad. En el precio lo cambiamos de número a número decimal fijo.

ID_Producto	Nombre_Producto	Cantidad	Precio_Unitario	Fecha_Ingreso
1	Teclado	47	669,54	miércoles, 1 de junio de 2022
2	Laptop	42	325,79	jueves, 2 de junio de 2022
3	Impresora	29	516,47	viernes, 3 de junio de 2022
4	Teclado	24	210,03	sábado, 4 de junio de 2022
5	Mouse	22	114,99	domingo, 5 de junio de 2022
6	Mouse	46	838,36	lunes, 6 de junio de 2022
7	Laptop	37	520,16	martes, 7 de junio de 2022
8	Cámara	8	652,74	miércoles, 8 de junio de 2022
9	Laptop	4	82,43	jueves, 9 de junio de 2022
10	Impresora	10	772,89	viernes, 10 de junio de 2022
11	Mouse	34	125,02	sábado, 11 de junio de 2022
12	Mouse	17	407,14	domingo, 12 de junio de 2022
14	Cámara	5	439,3	martes, 14 de junio de 2022
15	Teléfono	44	800,79	miércoles, 15 de junio de 2022
16	Teclado	18	746,9	jueves, 16 de junio de 2022
18	Teclado	31	901,05	sábado, 18 de junio de 2022
20	Impresora	14	191,2	lunes, 20 de junio de 2022
21	Teclado	20	632,71	martes, 21 de junio de 2022
22	Teléfono	40	197,23	miércoles, 22 de junio de 2022
23	Teléfono	27	690,51	jueves, 23 de junio de 2022
24	Mouse	4	887,56	viernes, 24 de junio de 2022
25	Cámara	28	140,63	sábado, 25 de junio de 2022
26	Mouse	49	53,22	domingo, 26 de junio de 2022
27	Impresora	36	464,48	lunes, 27 de junio de 2022
28	Mouse	45	601,35	martes, 28 de junio de 2022
29	Teléfono	5	382,72	miércoles, 29 de junio de 2022
31	Mouse	7	832,21	viernes, 1 de julio de 2022
32	Mouse	26	135,46	sábado, 2 de julio de 2022
33	Impresora	10	515,74	domingo, 3 de julio de 2022

PARTE 2

Ejercicio de práctica 2. Una vez realizado la limpieza y el proceso deberán guardar el archivo en el repositorio, deberás anexar la captura de pantalla con el nombre del archivo.