

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

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	0	28,87	jueves, 26 de enero de 2023
27	María	P005	82,57	viernes, 27 de enero de 2023
28	María	0	70,53	sábado, 28 de enero de 2023
29	Lucía	P002	12,89	domingo, 29 de enero de 2023

CONCLUSIONES

Limpieza: Se reemplazaron valores faltantes en códigos de productos, se estandarizaron los precios y se uniformaron los nombres de clientes. **Transformación:** Se normalizaron precios y se agruparon las transacciones por cliente. **Análisis:** Se identificaron tendencias de ventas, productos más vendidos y perfiles de clientes. La limpieza mejoró la precisión de los datos, permitiendo mejor planificación de inventario, estrategias de marketing más efectivas y mayor satisfacción del cliente, todo esto contribuye a decisiones más informadas y estratégicas para la empresa.

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

Paso a paso para la limpieza

ID_Cliente	Nombre_Cliente	Categoria_Conсульта	Duracion_Llamada	Fecha_Interaccion
2	José	Soporte Técnico	\$ 24,36	<i>lunes, 2 de enero de 2023</i>
3	Carlos	Facturación	\$ 58,86	<i>martes, 3 de enero de 2023</i>
4	José	Facturación	\$ 33,57	<i>miércoles, 4 de enero de 2023</i>
5	Pedro	Facturación	\$ 18,28	<i>jueves, 5 de enero de 2023</i>
7	Pedro	Soporte Técnico	\$ 29	<i>sábado, 7 de enero de 2023</i>
8	Luis	Consulta General	\$ 39,77	<i>domingo, 8 de enero de 2023</i>
9	Lucía	Soporte Técnico	\$ 23,21	<i>lunes, 9 de enero de 2023</i>
10	Ana	Consulta General	\$ 3,2	<i>martes, 10 de enero de 2023</i>
11	María	Quejas	\$ 2,37	<i>miércoles, 11 de enero de 2023</i>
12	José	Soporte Técnico	\$ 49,46	<i>jueves, 12 de enero de 2023</i>
13	Isabel	Quejas	\$ 26,74	<i>viernes, 13 de enero de 2023</i>
14	Pedro	Soporte Técnico	\$ 4,82	<i>sábado, 14 de enero de 2023</i>
15	Isabel	Consulta General	\$ 42,04	<i>domingo, 15 de enero de 2023</i>
16	Luis	Quejas	\$ 20,38	<i>lunes, 16 de enero de 2023</i>
17	Carlos	Facturación	\$ 19,76	<i>martes, 17 de enero de 2023</i>
18	Isabel	Quejas	\$ 51,29	<i>miércoles, 18 de enero de 2023</i>
19	Ana	Facturación	\$ 10,43	<i>jueves, 19 de enero de 2023</i>
20	María	Sugerencias	\$ 31,22	<i>viernes, 20 de enero de 2023</i>
21	Luis	Soporte Técnico	\$ 21,74	<i>sábado, 21 de enero de 2023</i>
22	Pedro	Quejas	\$ 48,28	<i>domingo, 22 de enero de 2023</i>
23	Isabel	Soporte Técnico	\$ 28,45	<i>lunes, 23 de enero de 2023</i>
24	Pedro	Soporte Técnico	\$ 22,27	<i>martes, 24 de enero de 2023</i>

CONCLUSIONES

Limpieza: Se estandarizaron los nombres de clientes, se completaron los valores faltantes en la categoría de consulta y se corrigieron los errores de formato en las duraciones de llamadas. **Transformación:** Se normalizaron los datos de duración de llamadas, asegurando consistencia en las unidades de tiempo y se categorizó cada consulta de manera uniforme. **Análisis:** Se identificaron patrones de consulta recurrentes, tiempos promedio de resolución por categoría y días con mayor volumen de llamadas. La limpieza mejoró significativamente la calidad de los datos, permitiendo identificar áreas con alta demanda y tiempos de resolución eficientes. Esto lleva a una mejor planificación de recursos, mejoras en la formación del personal y una mayor satisfacción del cliente, asegurando un servicio más eficiente y efectivo.

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".

El objetivo es optimizar la calidad de los datos de inventario mediante la corrección de errores comunes, asegurando que los registros sean consistentes, completos y adecuados para análisis de inventarios y toma de decisiones estratégicas en la tienda.

Data: data_inventario_para_limpieza 3

Paso a paso para la limpieza

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
13	Mouse	23	\$ 0	lunes, 13 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
17	Teléfono	10	\$ 0	viernes, 17 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

CONCLUSIONES

La limpieza y estandarización mejoraron la calidad de los datos, permitiendo un análisis más preciso y efectivo. Esto facilitó la toma de decisiones estratégicas, mejoró la gestión de inventarios y redujo costos, asegurando una reposición de stock más eficiente y alineada con la demanda del mercado.

