



Curso Power BI, Visualización de Datos e Inteligencia de Negocios

Módulo 2: Modelado de datos y transformación con Power Query

Sesión 1



Objetivos de hoy:

Al finalizar la sesión de hoy aprenderás...

- Que es la limpieza de datos
- Sobre calidad de datos y errores comunes
- Criterios en la calidad de datos
- Cuales son las operaciones básicas de limpieza.
- Y realizaras un ejercicio básico de limpieza de datos

Introducción a la limpieza de datos

- La limpieza de datos es el proceso de **identificar** y **corregir errores**, inconsistencias o valores atípicos en los conjuntos de datos.
- Es una etapa crítica antes de realizar análisis o visualizaciones, ya que garantiza la **calidad** y **confiabilidad de los resultados**.





Calidad de los datos y errores comunes

Los datos crudos generalmente contienen errores como :

1

Valores
nulos o
vacíos

2

Duplicados

3

Tipos de
datos
incorrectos

4

Inconsistencias
en nombres o
formatos

3

Datos fuera
de rango



Tipos comunes de errores en los datos

- **Valores nulos o vacíos:** Campos sin contenido pueden alterar cálculos o promedios.
- **Tipos de datos incorrectos:** Una fecha como texto impide filtros cronológicos.
- **Duplicados:** El mismo cliente o transacción se repite dos veces.
- **Inconsistencias en formato:** “Bogotá”, “bogota” y “BOGOTA” son tratados como distintos.

Ejemplo práctico:

Una base de datos de ventas tiene montos vacíos, fechas mal escritas y productos duplicados con distintas mayúsculas.



Calidad de los datos

Criterios

Precisión

Exactitud de los datos registrados.

Compleitud

Que todos los campos estén llenos.

Consistencia

Que no haya contradicciones entre columnas.

Validez

Que los datos estén dentro de los rangos esperados.



Impacto:

Una baja calidad de datos puede generar reportes incorrectos y decisiones erradas.

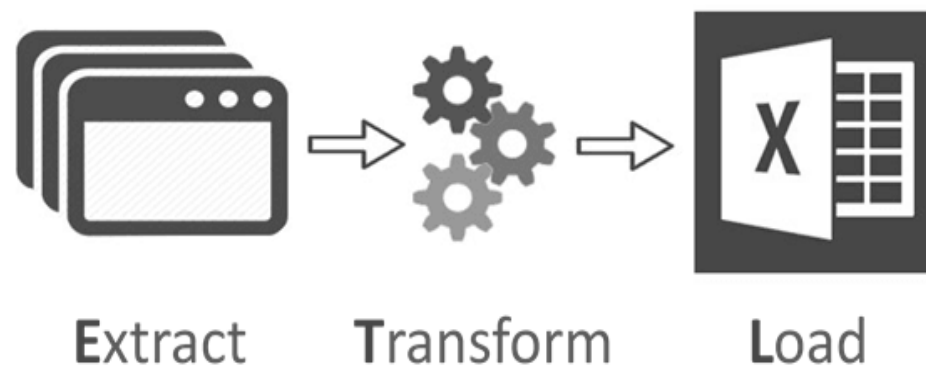


Power Query en Power BI

Es el editor de transformación de datos dentro de Power BI. Permite importar, limpiar, combinar y transformar datos sin afectar la fuente original.

Ventajas:

1. Registro de cada paso aplicado.
2. Automatización del proceso.
3. Vista previa instantánea.





Operaciones básicas de limpieza

1. Eliminar columnas y filas innecesarias.

Aplicaciones comunes

- Quitar columnas vacías o que no aportan al análisis.
- Eliminar filas con encabezados repetidos, comentarios o totales.

Herramientas en Power Query

- Botón “Quitar columnas”.
- Filtro de filas por condiciones.



Operaciones básicas de limpieza

2. Reemplazar y rellenar valores

¿Por qué reemplazar?

Evitar que valores nulos generen errores en visualizaciones o KPIs.

Reemplazar valores

Cambiar 'null' por "Sin dato" o 0.

Rellenar hacia abajo/arriba

Copiar valor de fila superior o inferior.



Operaciones básicas de limpieza

3. Modificar tipos de datos



Importancia:

Power BI necesita saber si una columna es texto, número o fecha para aplicar filtros y visualizaciones correctas..

Importancia:

Power BI necesita saber si una columna es texto, número o fecha para aplicar filtros y visualizaciones correctas.

Operaciones básicas de limpieza

4. Dividir y combinar columnas

Situaciones comunes:

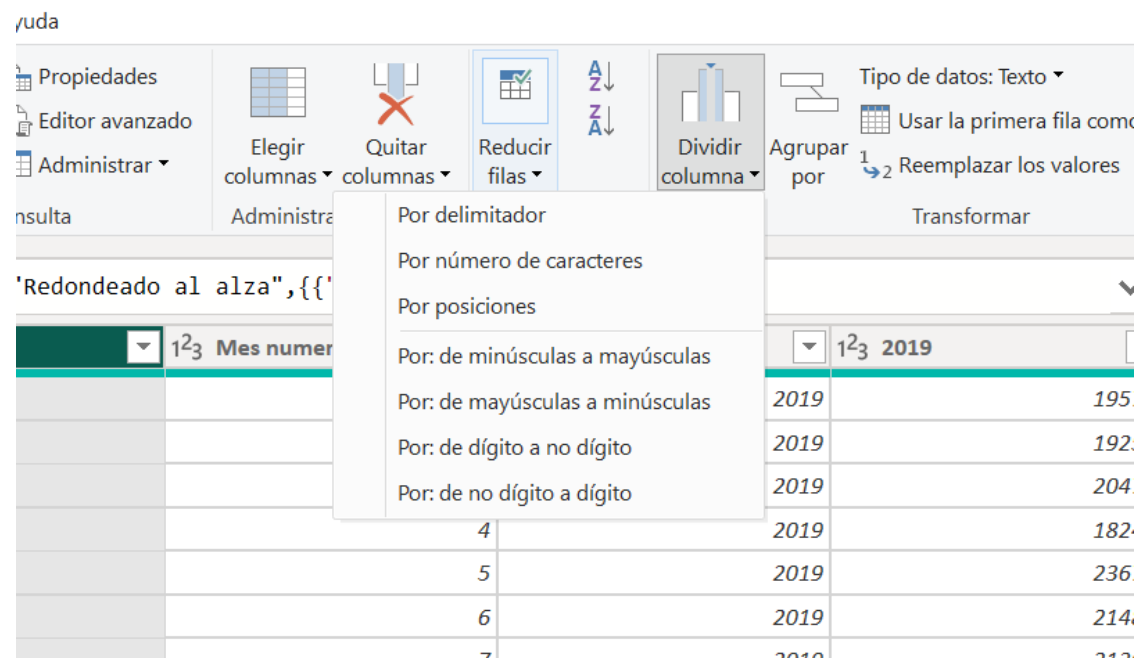
“Nombre completo” necesita separarse en “Nombre” y “Apellido”.

“Año-Mes” debe dividirse para usar jerarquías temporales.

Herramientas:

“Dividir por delimitador”.

“Combinar columnas” con separadores personalizados.





Operaciones básicas de limpieza

5. Quitar espacios y transformar texto

Problemas típicos:

- Espacios al inicio o final del texto impiden coincidencias exactas.
- Inconsistencia en mayúsculas/minúsculas complica agrupaciones.



Soluciones:

“Transformar texto” → Mayúsculas, minúsculas, quitar espacios.

Operaciones básicas de limpieza

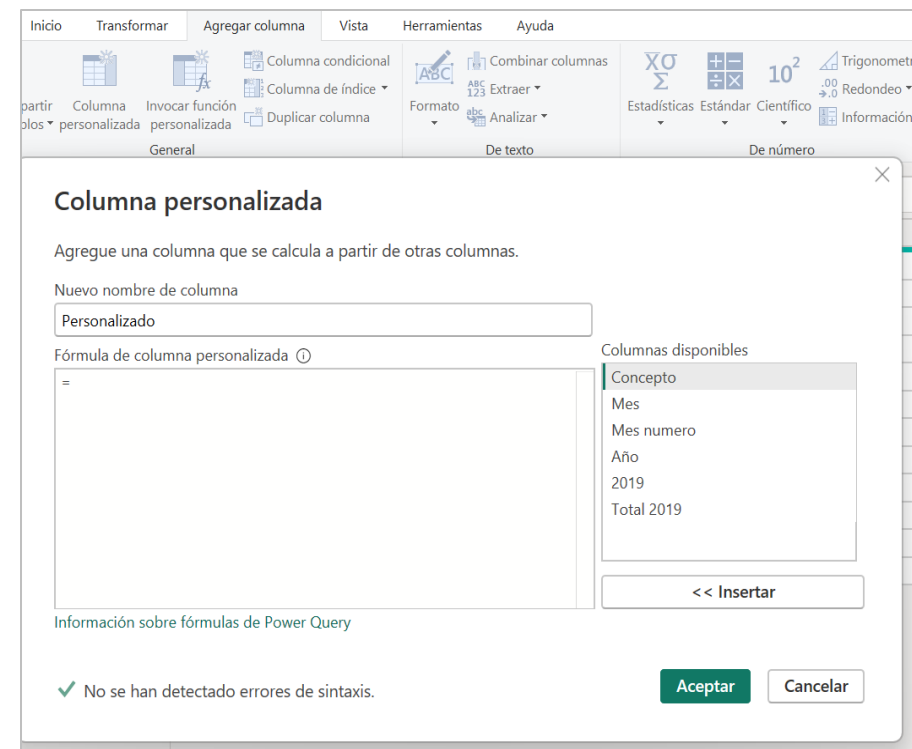
6. Columnas personalizadas

¿Qué es una columna personalizada?

Es una columna nueva creada con una fórmula, basada en el contenido de otras columnas.

Ejemplos:

- Clasificación de clientes por monto comprado
- Concatenación de nombre y apellido.
- Cálculo del margen: $(\text{Precio Venta} - \text{Costo}) / \text{Costo}$



Operaciones básicas de limpieza

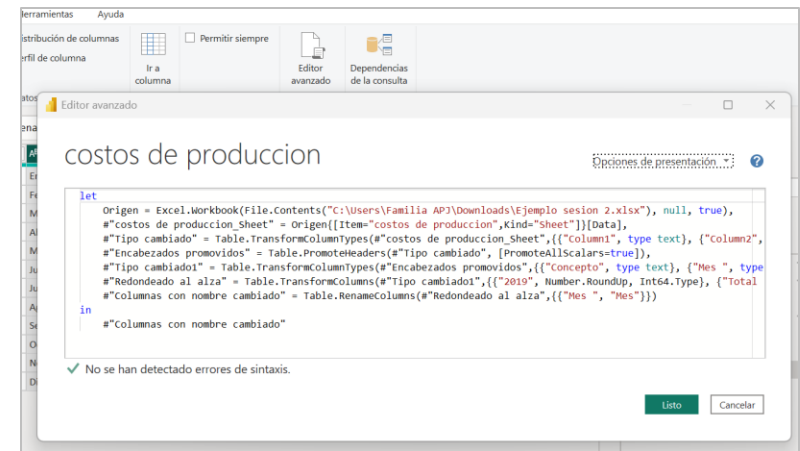
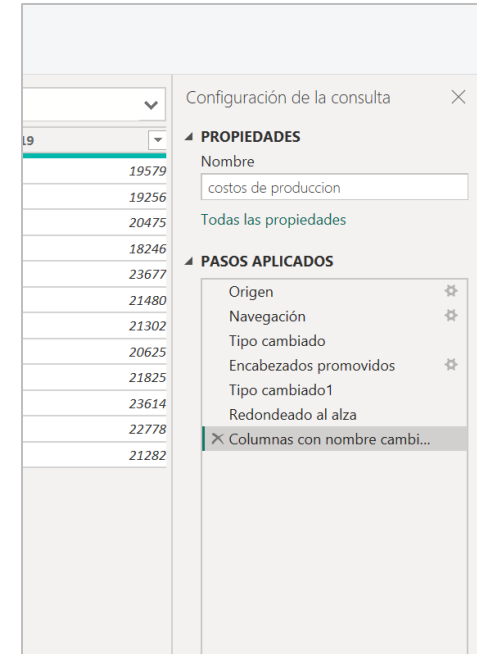
7. Pasos aplicados y editor avanzado

Ventajas del historial de pasos:

- Puedes ver, modificar o eliminar cualquier transformación.
- El orden de los pasos afecta el resultado final.

Editor avanzado:

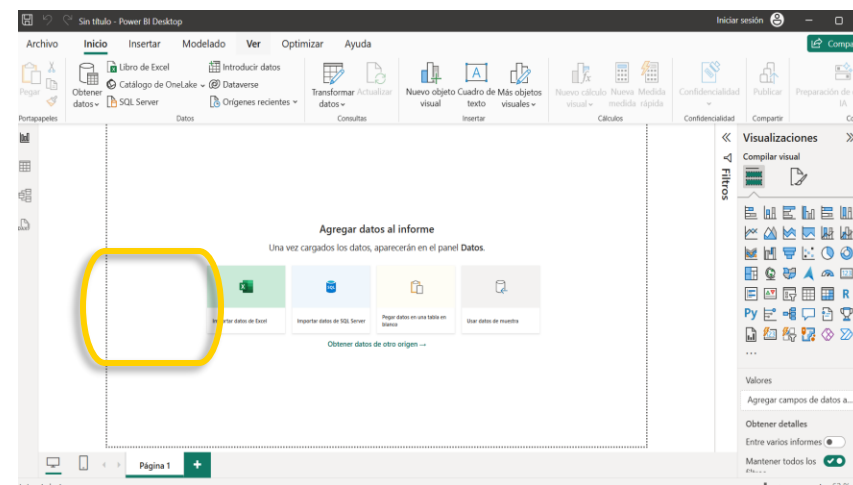
Muestra el código M detrás de las transformaciones. Útil para copiar, documentar o ajustar procesos complejos.



Como importar...

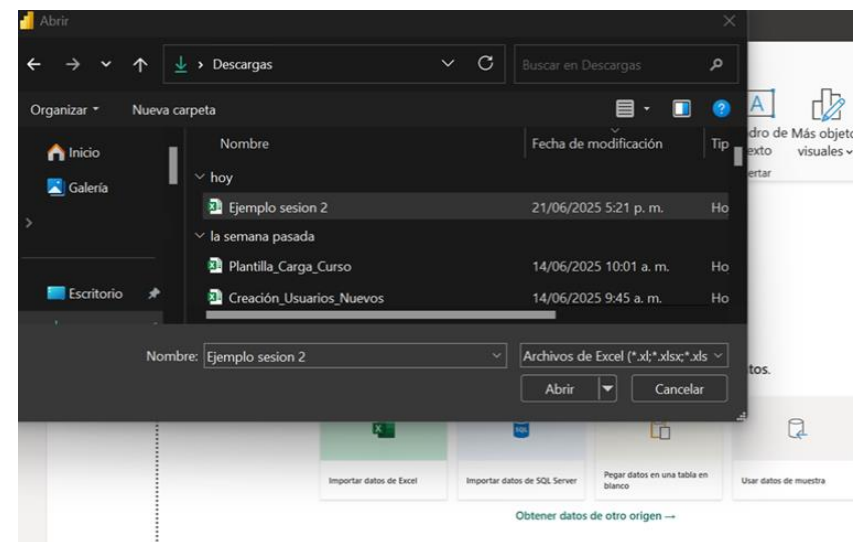
1

Seleccionar agregar datos al informe de una base excel



2

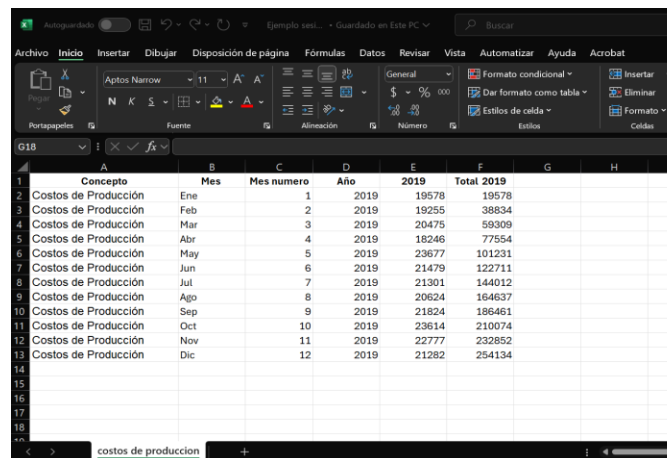
Ejemplo : Importar base de excel



Iniciemos...

3

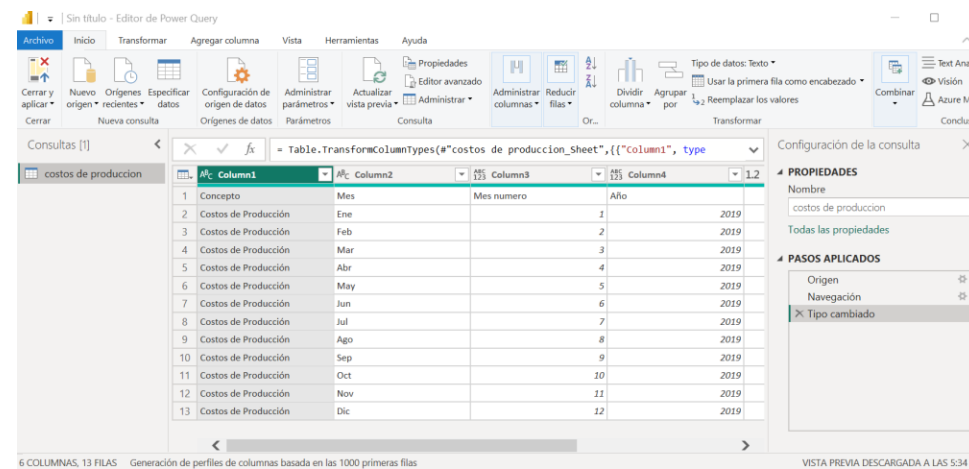
Ejemplo: Base de datos de costos con mes, año, datos de costos total



| Concepto | Mes | Mes numero | Año | 2019 | Total 2019 |
|----------------------|-----|------------|------|-------|------------|
| Costos de Producción | Ene | 1 | 2019 | 19578 | 19578 |
| Costos de Producción | Feb | 2 | 2019 | 19255 | 38834 |
| Costos de Producción | Mar | 3 | 2019 | 20475 | 59309 |
| Costos de Producción | Abr | 4 | 2019 | 18246 | 77554 |
| Costos de Producción | May | 5 | 2019 | 23677 | 101231 |
| Costos de Producción | Jun | 6 | 2019 | 21479 | 122711 |
| Costos de Producción | Jul | 7 | 2019 | 21361 | 144012 |
| Costos de Producción | Ago | 8 | 2019 | 20624 | 164637 |
| Costos de Producción | Sep | 9 | 2019 | 21824 | 186461 |
| Costos de Producción | Oct | 10 | 2019 | 23614 | 210074 |
| Costos de Producción | Nov | 11 | 2019 | 22777 | 232852 |
| Costos de Producción | Dic | 12 | 2019 | 21282 | 254134 |

4

En Power BI podemos dar click en transformar datos para la limpieza a la base como queremos que quede



| Concepto | Mes | Mes numero | Año | 2019 |
|-------------------------|-----|------------|-----|------|
| 1 Concepto | Mes | | | |
| 2 Costos de Producción | Ene | | | 2019 |
| 3 Costos de Producción | Feb | | | 2019 |
| 4 Costos de Producción | Mar | | | 2019 |
| 5 Costos de Producción | Abr | | | 2019 |
| 6 Costos de Producción | May | | | 2019 |
| 7 Costos de Producción | Jun | | | 2019 |
| 8 Costos de Producción | Jul | | | 2019 |
| 9 Costos de Producción | Ago | | | 2019 |
| 10 Costos de Producción | Sep | | | 2019 |
| 11 Costos de Producción | Oct | | | 2019 |
| 12 Costos de Producción | Nov | | | 2019 |
| 13 Costos de Producción | Dic | | | 2019 |



Pongamos en práctica: Limpieza de datos...

De manera individual importaremos un Excel con errores intencionales para ello vamos a:

1. Guardar archivo Excel con errores (columnas vacías entre ellas, datos con formato texto, encabezados con nombres erróneos, texto con formato número) en equipo.
2. Importar archivo a Power BI **Base Productos Modulo 2 Sesion 1 Ejercicio práctico.xlsx**
3. En el archivo iremos a **transformar datos** y vamos a: Eliminar columnas vacías , renombrar encabezados, reemplazar valores nulos y cambiar tipos de datos
4. **Tomar pantallazos y subir al foro del curso como evidencia del ejercicio práctico.**

Ver: Manual ejercicio práctico Modulo 2 Sesión 1 -Pongamos en práctica

