

# Limpiar datos con PBI

<b>Objetivo</b>	<b>1</b>
<b>Instrucciones</b>	<b>1</b>
Importar los archivos CSV	1
Limpieza de clientes_dirty.csv	2
Problemas detectados	2
Pasos de limpieza	2
Limpieza de pedidos_dirty.csv	2
Problemas detectados	2
Pasos de limpieza	2
Limpieza de productos_dirty.csv y detalle_pedidos_dirty.csv	3
<b>Resumen de herramientas de Power Query usadas</b>	<b>3</b>
<b>Entregables</b>	<b>4</b>

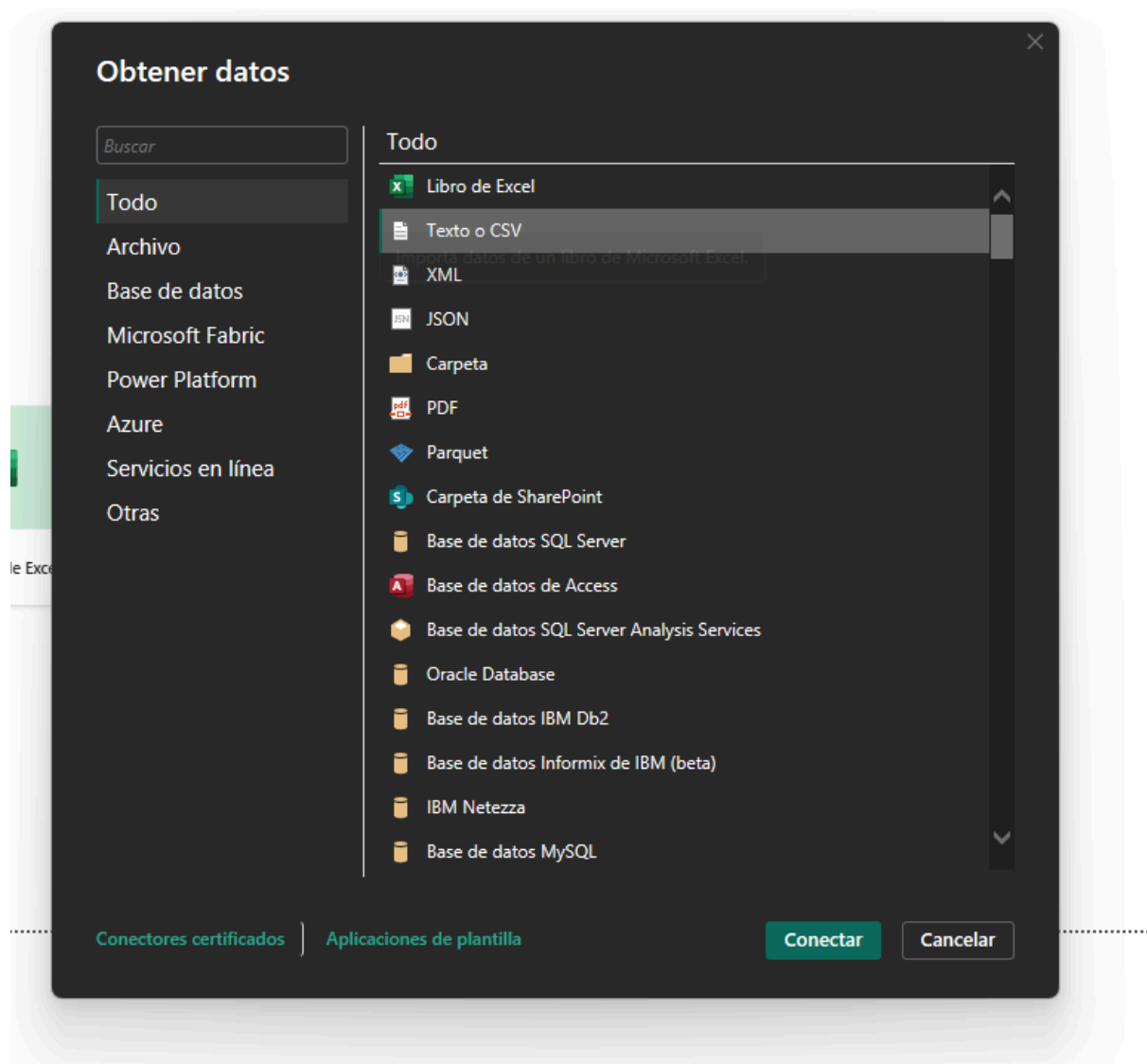
## Objetivo

Aprender a **limpiar y preparar datos sucios** con Power BI Desktop.

## Instrucciones

### Importar los archivos CSV

1. **Descomprime** el archivo zip adjunto *PowerBI\_dirty.zip*.
2. Abre **Power BI Desktop**.
3. Haz clic en **Obtener datos > Texto/CSV**.



4. Selecciona los archivos:
  - clientes\_dirty.csv



- pedidos\_dirty.csv
  - productos\_dirty.csv
  - detalle\_pedidos\_dirty.csv
5. Haz clic en **Cargar** o **Transformar datos** para ir al **Editor de Power Query**.

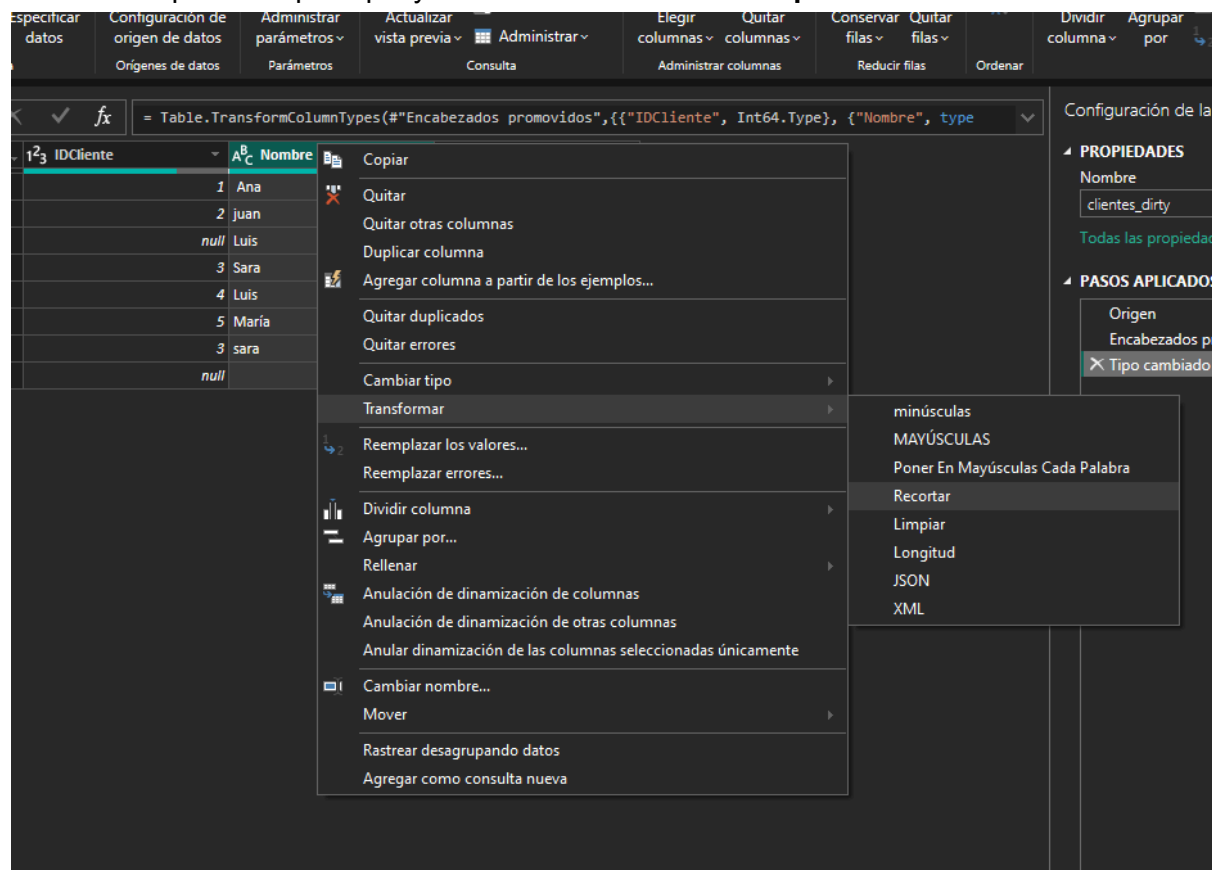
# Limpieza de clientes\_dirty.csv

## Problemas detectados

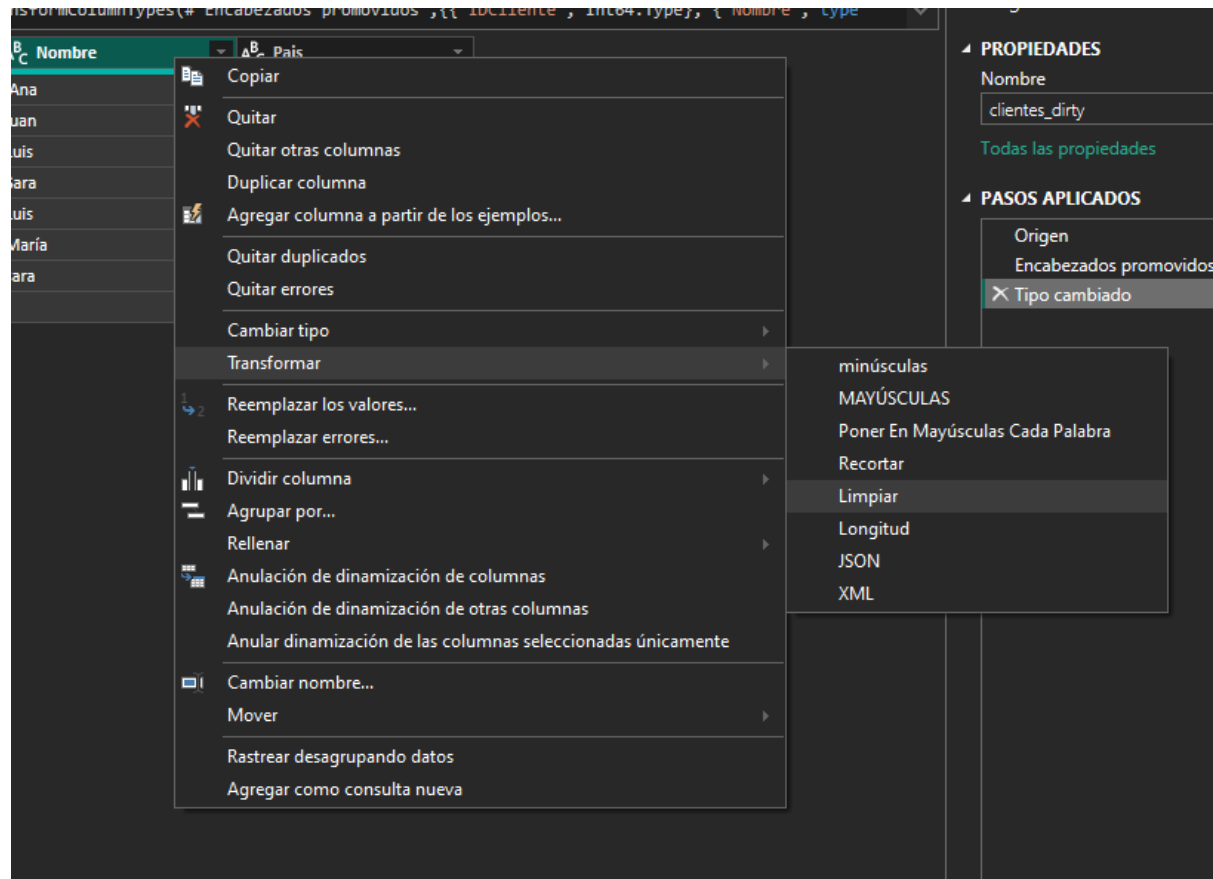
- Espacios en los nombres.
- Mayúsculas y minúsculas inconsistentes.
- Países con espacios y acentos inconsistentes.

## Pasos de limpieza

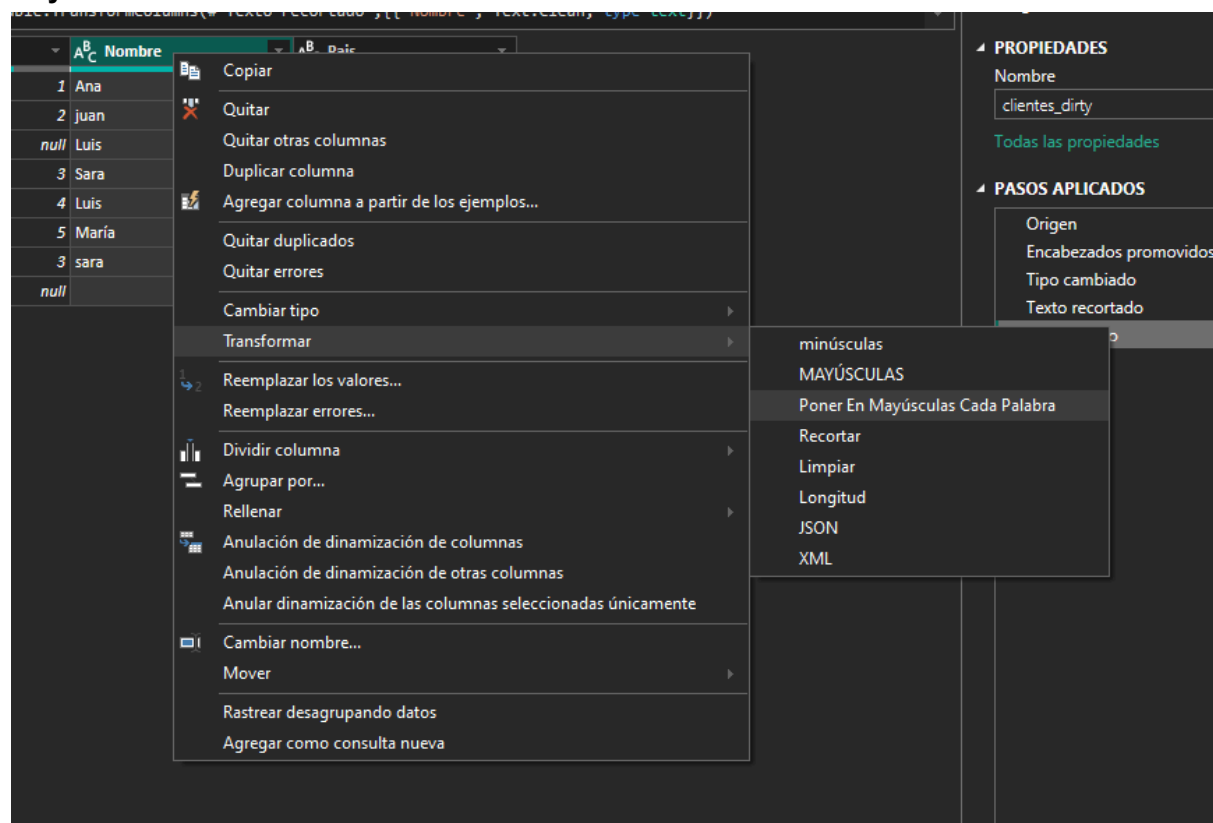
1. Selecciona la columna *Nombre*.
  - Quitamos espacios al principio y al final: **Transformar > Limpiar > Recortar**.



- Quitamos caracteres no imprimibles: **Transformar > Limpiar > Limpiar.**

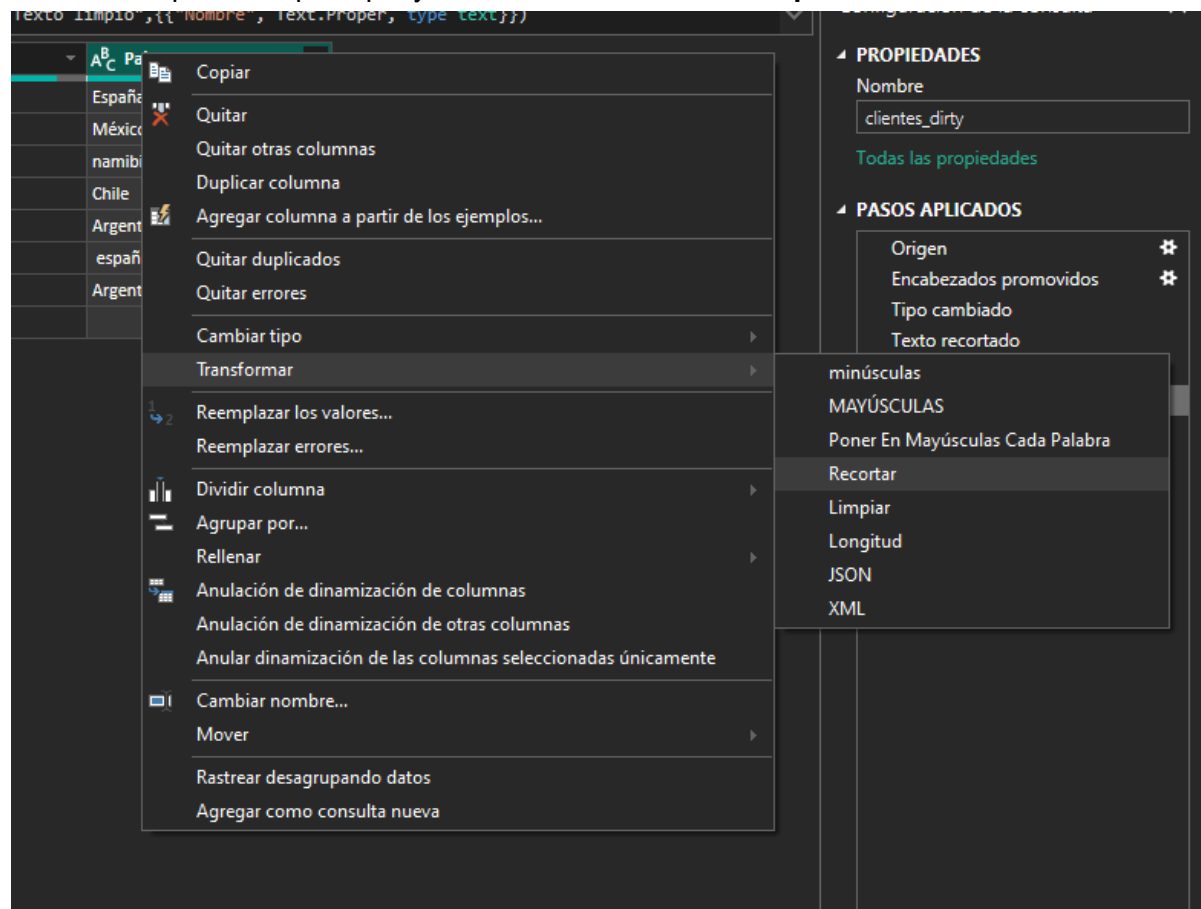


- Ponemos en mayúsculas cada palabra: **Transformar > Formato > Poner En Mayúsculas Cada Palabra.**

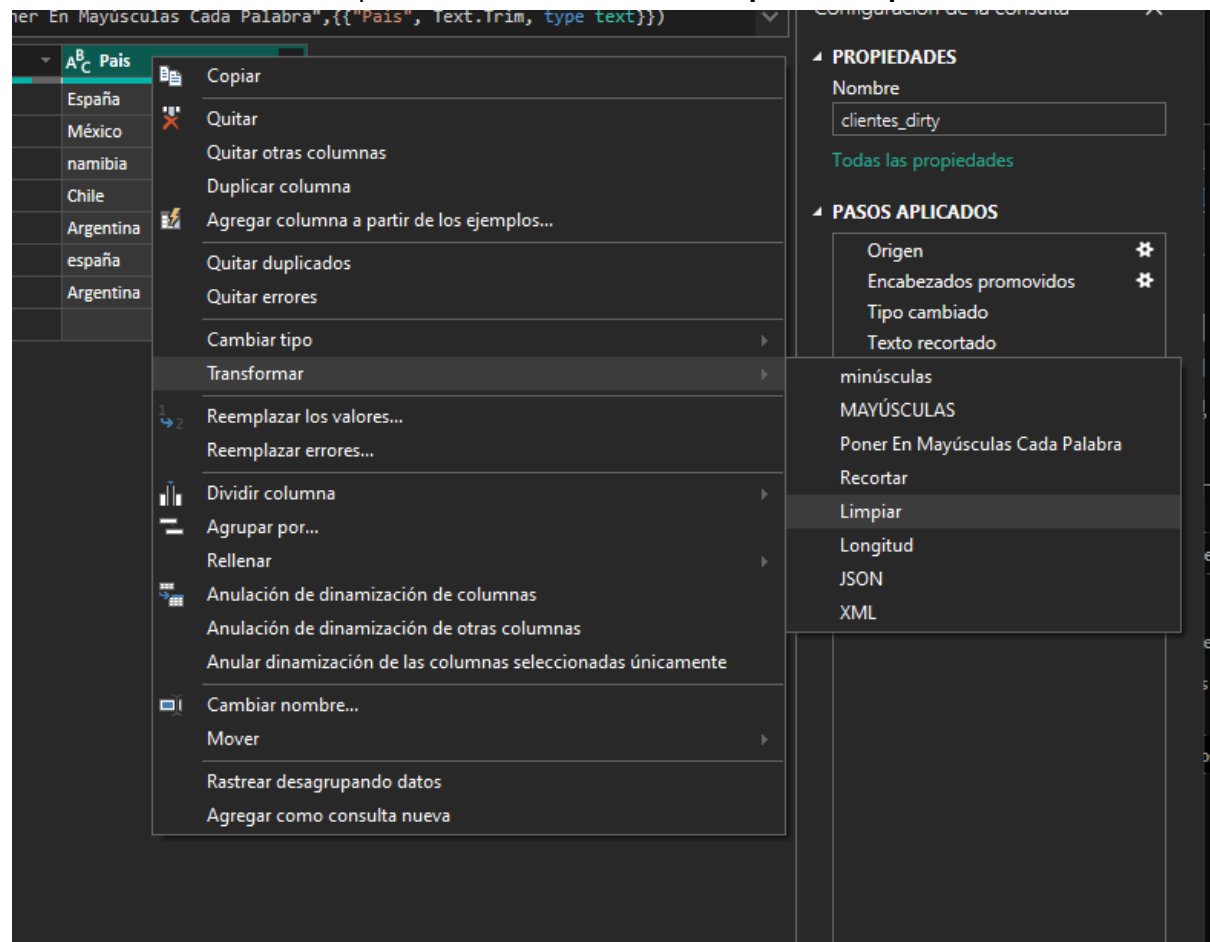


2. Selecciona la columna *Pais*.

- Quitamos espacios al principio y al final: **Transformar > Limpiar > Recortar**.



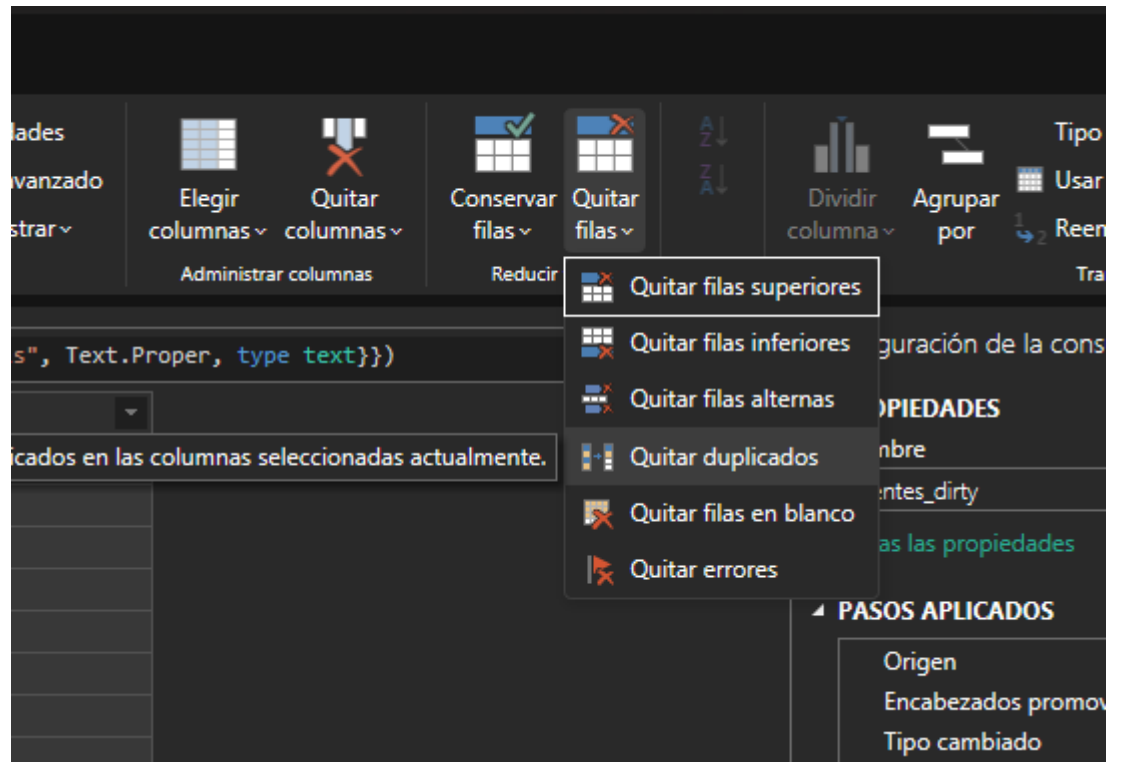
- Quitamos caracteres no imprimibles: **Transformar > Limpiar > Limpiar.**



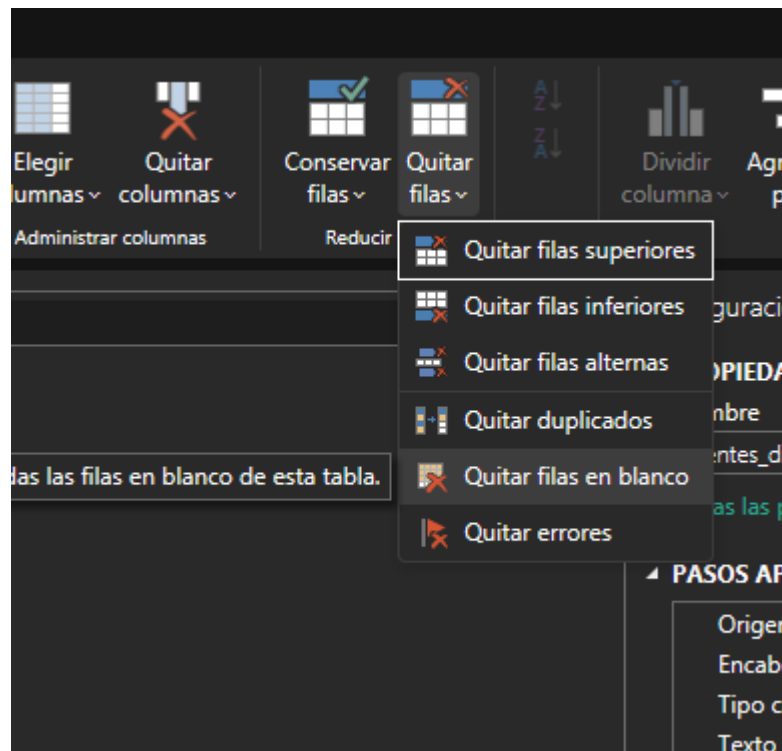
- Ponemos en mayúsculas cada palabra: **Transformar > Formato > Poner En Mayúsculas Cada Palabra.**

3. Selecciona todas las columnas.

- Eliminamos las filas duplicadas: **Inicio > Quitar filas > Quitar duplicados**.

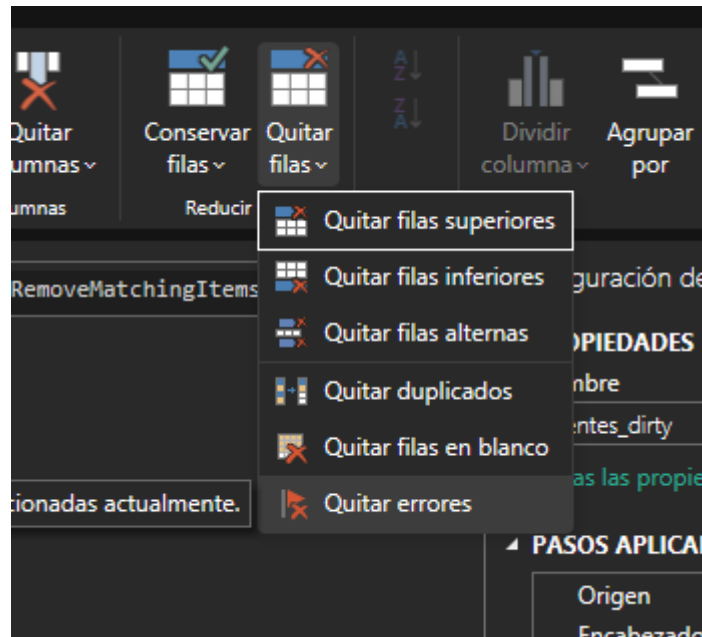


- Eliminamos las filas vacías: **Inicio > Quitar filas > Quitar filas en blanco**.



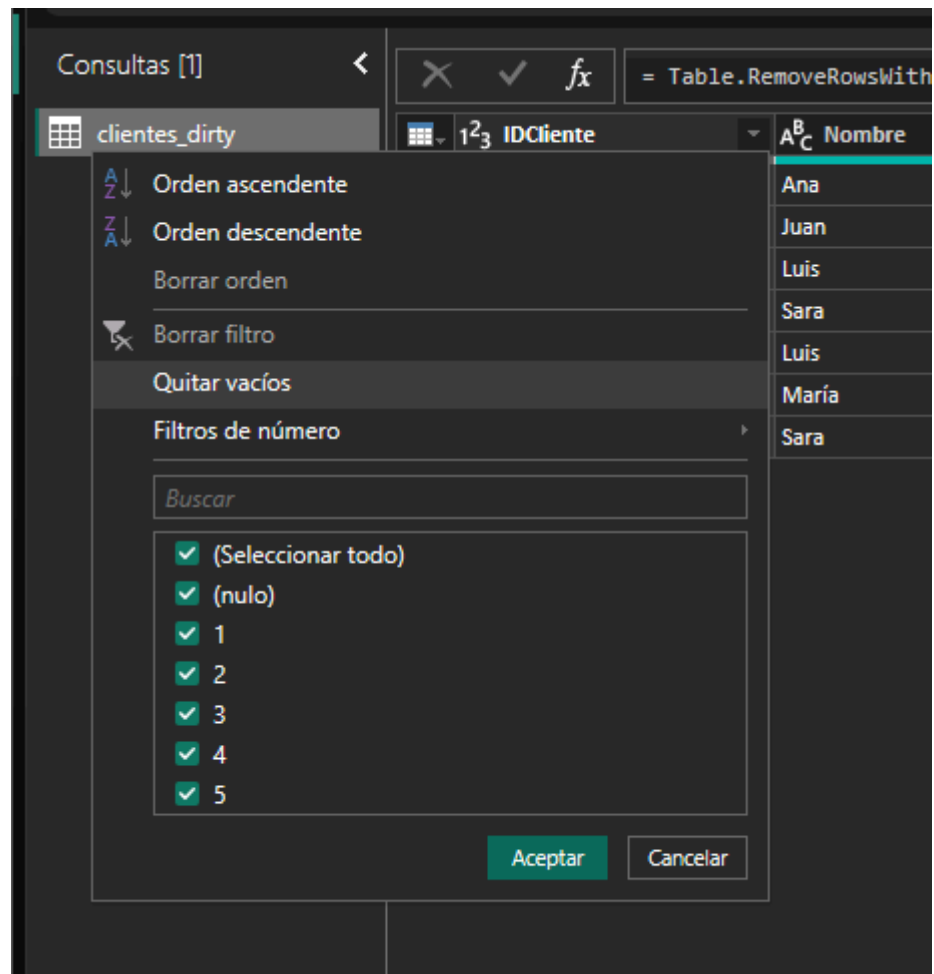


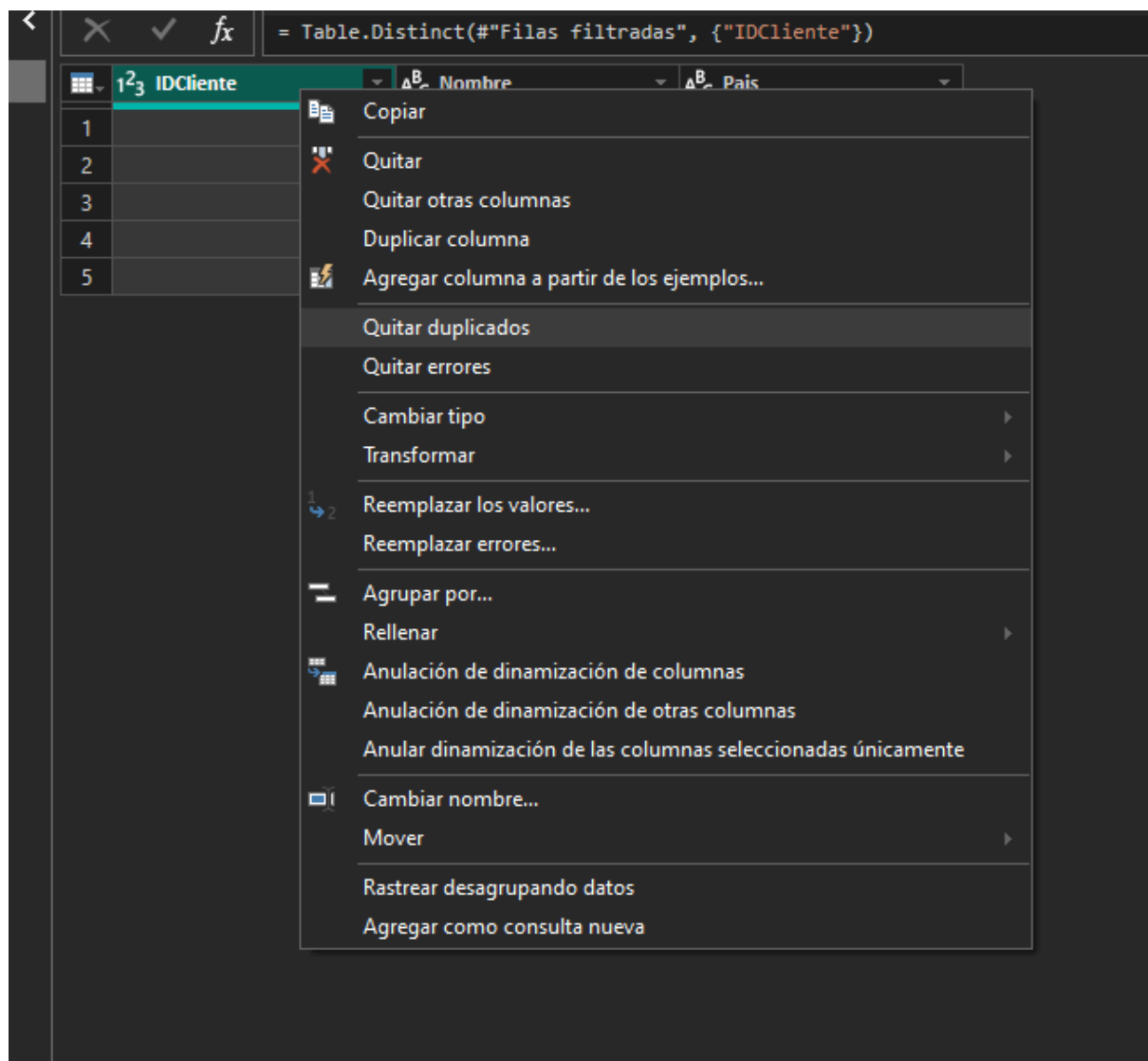
- Eliminamos las filas con errores: **Inicio > Quitar filas > Quitar errores**.



4. Selecciona la columna *IDCliente*.

- Selecciona el filtro: **Quitar vacíos**.





## Limpieza de pedidos\_dirty.csv

### Problemas detectados

- Fechas en formatos diferentes (10/01/2023, 12-03-2023, etc.).
- Importe con algunos decimales y otros enteros.

### Pasos de limpieza

Selecciona la columna *Fecha*.

- Haz clic en **Transformar > Tipo de datos > Fecha**.
- Power Query convertirá automáticamente distintos formatos de fecha.

Selecciona la columna *Importe*.

- Haz clic en **Transformar > Tipo de datos > Número decimal**.

Verifica que IDCliente tenga el mismo tipo de dato que en clientes\_dirty.csv (Número entero).

- Elimina filas duplicadas, vacías y con errores por si en el futuro las hubiera.



## Limpieza de productos\_dirty.csv y detalle\_pedidos\_dirty.csv

*productos\_dirty.csv* está casi limpio:

- Asegúrate de que IDProducto sea Número entero.
- Precio como Número decimal.
- Elimina filas duplicadas, vacías y con errores por si en el futuro las hubiera.

*detalle\_pedidos\_dirty.csv*:

- IDPedido y IDProducto como Número entero.
- Cantidad como Número entero.
- Elimina filas duplicadas, vacías y con errores por si en el futuro las hubiera.

## Resumen de herramientas de Power Query usadas

Acción	Ubicación
Quitar espacios	Transformar > Limpiar > Recortar
Quitar caracteres no imprimibles	Transformar > Limpiar > Limpiar
Capitalizar	Transformar > Formato > Poner En Mayúsculas Cada Palabra
Cambiar tipo de dato	Transformar > Tipo de datos
Quitar filas duplicados	Inicio > Quitar filas > Quitar duplicados
Quitar filas vacías	Inicio > Quitar filas > Quitar duplicados
Quitar filas con errores	Inicio > Quitar filas > Quitar errores
Quitar columnas vacías o con errores	Filtrar columnas y quitar vacíos
Separar columnas	Transformar > Columna > Dividir columna (si aplica)

# Entregables

- Añadir como comentarios la tarea un enlace a Github, de acceso público, a un repositorio que contenga el archivo **Power BI: nombreyapellidos-A3-2-Limpiar datos con PBI.pbix**.
- Archivo **PDF: nombreyapellidos-A3-2-Limpiar datos con PBI.pdf**. Hazlo a partir de este archivo añadiendo, en el apartado correspondiente:
  - Capturas de pantalla de ~~autoría inequívoca~~.
  - Comentarios y observaciones.