



Sesión 01

19-21/05/2024

Power BI

Tema del día

Inicio	Duración	Fin	Actividad
8:00	0:30	8:30	Presentación
8:30	1:15	9:45	Importar multiples orígenes/Group by/Union
9:45	0:15	10:00	Break
10:00	0:45	10:45	Funciones personalizadas
10:45	1:15	12:00	Normalizar datos (casos 1/2/3)

Importar
múltiples
orígenes

CVS

Excel

Página web

PDF

Otros...

Cargar archivo desde repositorio en la nube

- OneDrive
- GoogleDrive
- ...

Group by (Agrupamiento)

Funcionalidad que permite consolidar datos en grupos basados en uno o más campos específicos. Esto se utiliza para calcular sumas, promedios, conteos y otras agregaciones dentro de cada grupo.

- **Pasos básicos para usar Group By en Power BI:**
 - 1. Seleccionar los datos:** Identificar la columna o columnas por las cuales se desea agrupar.
 - 2. Aplicar el agrupamiento:** Utilizar la herramienta de Power Query o DAX (Data Analysis Expressions) para crear la agregación deseada.
 - 3. Realizar cálculos agregados:** Dentro de cada grupo, realizar operaciones como sumas, promedios, conteos, etc.

Group by (Agrupamiento)

- **Ejemplo:** Si tienes una tabla de ventas con columnas para la fecha, el producto y la cantidad vendida, puedes usar Group By para sumar la cantidad vendida por cada producto.
- **Herramientas en Power BI para Group By:**
 - **Power Query:** Ofrece una interfaz gráfica para agrupar datos.
 - **DAX:** Permite agrupar datos mediante expresiones y fórmulas más avanzadas.

Union (Unión)

Se utiliza para combinar filas de dos o más tablas en una sola tabla, conservando las columnas de ambas. Las tablas unidas deben tener el mismo número de columnas y tipos de datos correspondientes en cada columna.

- **Pasos para usar Unión en Power BI:**
 - **Seleccionar las tablas:** Identificar las tablas que se desean unir.
 - **Aplicar la función Union:** Utilizar DAX (Data Analysis Expressions) para crear una nueva tabla que contenga las filas de ambas tablas.

Union (Unión)

- **Sintaxis básica en DAX:**

```
UNION (Tabla1, Tabla2)
```

- **Ejemplo:** Si tienes dos tablas de ventas de diferentes regiones, puedes usar la función Union para crear una tabla combinada que incluya las ventas de ambas regiones.

Funciones personalizadas

Permiten a los usuarios crear sus propias funciones reutilizables dentro del entorno de Power Query. Estas funciones son útiles para simplificar y automatizar tareas repetitivas o complejas durante el proceso de transformación de datos.

- **Pasos para crear una función personalizada en Power BI:**
 - **Definir la función:** Crear la lógica de la función usando el lenguaje M.
 - **Guardar la función:** Convertir una consulta en una función que se puede llamar desde otras consultas.
 - **Invocar la función:** Aplicar la función personalizada a una o más tablas para realizar operaciones específicas.

Funciones personalizadas

- **Ejemplo:** Si necesitas calcular el impuesto sobre las ventas en múltiples tablas, puedes crear una función personalizada que acepte la cantidad de ventas como parámetro y devuelva el impuesto calculado.

```
let
  CalculateTax=(SalesAmount as number) as number =>
  let
    TaxRate = 0.1,
    TaxAmount = SalesAmount * TaxRate
  in
    TaxAmount
in
  CalculateTax
```

Normalizar datos

La **normalización de datos** en Power BI se refiere a la práctica de organizar los datos de manera que minimicen la redundancia y dependencias, mejorando así la consistencia y la eficiencia en el análisis.

Se tienen tres casos principales de normalización de datos...

Caso 1: Normalización básica de datos

Consiste en la limpieza y estructuración de datos para eliminar duplicados y asegurar que cada dato tenga un único punto de referencia. Este proceso generalmente incluye:

- **Eliminar duplicados:** Garantizar que cada registro sea único.
- **Estandarizar formatos:** Uniformar formatos de fechas, números, y textos.
- **Separar datos combinados:** Dividir columnas que contienen múltiples tipos de información en columnas separadas.

Caso 2: Normalización de datos no homogéneos

Se enfoca en transformar datos de diferentes fuentes con estructuras distintas en un formato homogéneo. Este proceso incluye:

- **Mapeo de campos:** Alinear campos de diferentes tablas que representan la misma información pero tienen nombres diferentes.
- **Unificación de unidades:** Convertir distintas unidades de medida a una unidad estándar.
- **Estandarización de categorías:** Uniformar categorías o clasificaciones que pueden variar entre fuentes.

Caso 3: Modelamiento y normalización avanzada

Implica la creación de modelos de datos más complejos que soporten análisis robustos, incluyendo:

- **Desnormalización controlada:** Combinar tablas cuando se requiere mejorar el rendimiento de consultas específicas.
- **Creación de relaciones:** Definir relaciones entre tablas para permitir análisis multidimensionales.
- **Optimización de modelos:** Ajustar el modelo de datos para mejorar la eficiencia y velocidad de los cálculos.

Normalización vs Desnormalización

1	Implementación	La normalización se utiliza para eliminar datos redundantes de la base de datos y para almacenar datos no redundantes y coherentes en ella.	La desnormalización se utiliza para combinar varios datos de tabla en uno para que se pueda consultar rápidamente.
2	centro de atención	La normalización se centra principalmente en borrar la base de datos de los datos no utilizados y reducir la redundancia y la incoherencia de los datos.	La desnormalización, por otro lado, se centra en lograr una ejecución más rápida de las consultas mediante la introducción de redundancia.
3	Número de tablas	Durante la normalización, los datos se reducen, por lo que se eliminan varias tablas de la base de datos, por lo que las tablas tienen un número menor.	Por otro lado, durante la desnormalización, los datos se integran en la misma base de datos y, por lo tanto, una serie de tablas para almacenar ese número de datos aumenta.
4	Consumo de memoria	La normalización utiliza memoria optimizada y, por lo tanto, un rendimiento más rápido.	Por otro lado, la desnormalización introduce algún tipo de desperdicio de memoria.
5	Integridad de los datos	La normalización mantiene la integridad de los datos, es decir, cualquier adición o eliminación de datos de la tabla no creará ninguna discrepancia en la relación de las tablas.	La desnormalización no mantiene ninguna integridad de datos.
6	Dónde usar	La normalización se utiliza generalmente cuando se realiza el número de operaciones de inserción/actualización/eliminación y las combinaciones de esas tablas no son costosas.	Por otro lado, la desnormalización se utiliza cuando las combinaciones son costosas y se ejecuta una consulta frecuente en las tablas.