

RETO 2 ITINERARIO BUSINESS ANALITYCS FUNDAC

DOCUMENTO DE ELABORACION

De acuerdo a las instrucciones para la elaboración de este trabajo, he optado por un dataset disponible en **Kaggle**

Características del dataset (disponible en el propio sitio web Kaggle)

Contexto

Ante la creciente demanda y la feroz competencia en el mercado, un gigante de las grandes superficies busca su conocimiento para comprender qué les funciona mejor. Quieren comprender qué productos, regiones, categorías y segmentos de clientes deberían dirigirse o evitar.

Incluso puedes llevar esto un paso más allá e intentar construir un modelo de regresión para predecir las ventas o las ganancias.

Vuélvete loco con el conjunto de datos, pero asegúrate también de proporcionar información comercial para mejorar.

Metadatos

ID de fila => ID único para cada fila.

ID de pedido => ID de pedido único para cada cliente.

Fecha de pedido => Fecha de pedido del producto.

Fecha de envío => Fecha de envío del producto.

Modo de envío => Modo de envío especificado por el cliente.

ID de cliente => ID único para identificar a cada cliente.

Nombre del cliente => Nombre del cliente.

Segmento => El segmento al que pertenece el cliente. País => País de residencia

del cliente. Ciudad

=> Ciudad de residencia del cliente. Estado => Estado de residencia del cliente

. Código postal => Código postal de cada cliente.

Región => Región a la que pertenece el cliente.

ID de producto => ID único del producto.

Categoría => Categoría del producto pedido.

Subcategoría => Subcategoría del producto pedido.

Nombre del producto => Nombre del producto

Ventas => Ventas del producto.

Cantidad => Cantidad del producto.

Descuento => Descuento proporcionado.

Beneficio => Ganancia/Pérdida incurrida.

Tras importar este archivo, se ha procedido a comprobar y depurar los datos

Se presentan solo los 1.000 primeros registros. Se cargan la totalidad.

Comprobamos la calidad de las columnas (*Inicio* → *Vista* → *Calidad de las columnas*): no existen errores en los datos importados

Inicio Transformar Agregar columna **Vista** Herramientas Ayuda

☐ Configuración de la consulta
 ☐ Monoespaciada
 ☐ Distribución de columnas
 ☐ Permitir siempre
 Ir a columna
 Editor avanzado
 Dependencias de la consulta

☒ Barra de fórmulas
 ☒ Mostrar espacio en blanco
 ☐ Perfil de columna
 Calidad de columnas

Diseño Vista previa de datos Parámetros Columnas Uso avanzado Dependencias

= Table.SelectRows("#Columnas con nombre cambiado", each true)

ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID
9	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710
10	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710
11	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710
12	CA-2014-115812	6/9/2014	6/14/2014	Standard Class	BH-11710
13	CA-2017-114412	4/15/2017	4/20/2017	Standard Class	AA-10480
14	CA-2016-161389	12/5/2016	12/10/2016	Standard Class	IM-15070
15	US-2015-118983	11/22/2015	11/26/2015	Standard Class	HP-14815
16	US-2015-118983	11/22/2015	11/26/2015	Standard Class	HP-14815

Se realiza la comprobación de la distribución de las columnas (*Inicio → Vistas → Distribución de las columnas*)

Inicio Transformar Agregar columna **Vista** Herramientas Ayuda

☐ Configuración de la consulta
 ☐ Monoespaciada
 ☒ Distribución de columnas
 ☐ Permitir siempre
 Ir a columna
 Editor avanzado
 Dependencias de la consulta

☒ Barra de fórmulas
 ☒ Mostrar espacio en blanco
 ☐ Perfil de columna
 Calidad de columnas

Diseño Vista previa de datos Parámetros Columnas Uso avanzado Dependencias

= Table.SelectRows("#Columnas con nombre cambiado", each true)

Category	Sub-Category	Product Name	Sales
Office Supplies	Binders	DAL Angie-view binders with locking rings by samisii	18304
Office Supplies	Appliances	Belkin F5C206VTEL 6 Outlet Surge	1149
Furniture	Tables	Chromcraft Rectangular Conference Tables	1706184
Technology	Phones	Konftel 250 Conference phone - Charcoal black	911424
Office Supplies	Paper	Xerox 1967	15552

Se cambian los nombres de las columnas al idioma castellano

Durante el proceso de ETL vamos a crear una columna calculada para obtener dos nuevas columnas, en las que se separe el nombre y el apellido de cada cliente

Transformar → Dividir columna → por delimitador

RETO2_FranciscoAlejandroCastillaDonis

Inicio **Transformar** Agregar columna Vista Herramientas Ayuda

Agrupar por Usar la primera fila como encabezado Contar filas Transponer Invertir filas

Tipo de datos: Texto Reemplazar los valores Anular dinamización de columnas
 Detectar tipo de datos Rellenar Mover
 Cambiar nombre Columna dinámica Convertir en lista

Divide los valores de la columna seleccionada según el delimitador especificado.

Dividir columna Formato Extraer Analizar

Consultas [5]

Sample - Superstore

ID Cliente	Nombre Cliente	Segmento
CG-12520	Claire Gute	Consumidor
CG-12520	Claire Gute	Consumidor
DV-13045	Darrin Van Huff	Corporativo
SO-20335	Sean O'Donnell	Consumidor
SO-20335	Sean O'Donnell	Consumidor
BH-11710	Brosina Hoffman	Consumidor
BH-11710	Brosina Hoffman	Consumidor

Por delimitador
 Por número de caracteres
 Por posiciones
 Por: de minúsculas a mayúsculas
 Por: de mayúsculas a minúsculas
 Por: de dígito a no dígito
 Por: de no dígito a dígito

Renombramos las dos nuevas columnas como Nombre y Apellido

Vamos a extraer de la columna Fecha Pedido, el año

Agregar columna → extraer → últimos caracteres → 4

RETO2_FranciscoAlejandroCastillaDonis

Inicio Transformar **Agregar columna** Vista Herramientas Ayuda

Columna a partir de los ejemplos Columna personalizada Invocar función personalizada

Columna condicional Columna de índice Combinar columnas Extraer

General

Consultas [5]

Sample - Superstore

ID Cliente	Nombre Cliente	Segmento
CG-12520	Claire Gute	Consumidor
CG-12520	Claire Gute	Consumidor
DV-13045	Darrin Van Huff	Corporativo
SO-20335	Sean O'Donnell	Consumidor
SO-20335	Sean O'Donnell	Consumidor
BH-11710	Brosina Hoffman	Consumidor
BH-11710	Brosina Hoffman	Consumidor

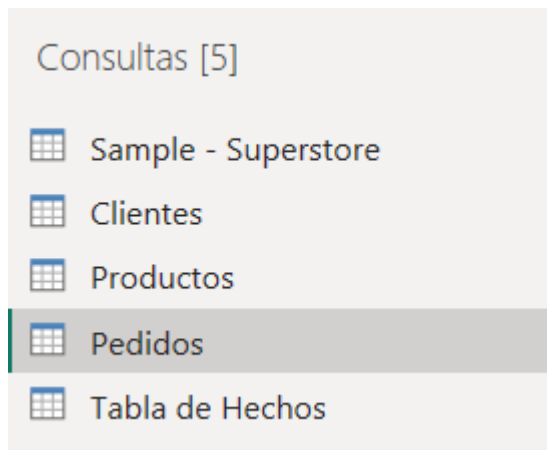
Insertar últimos caracteres

Escriba cuántos caracteres finales desea conservar.

Recuento

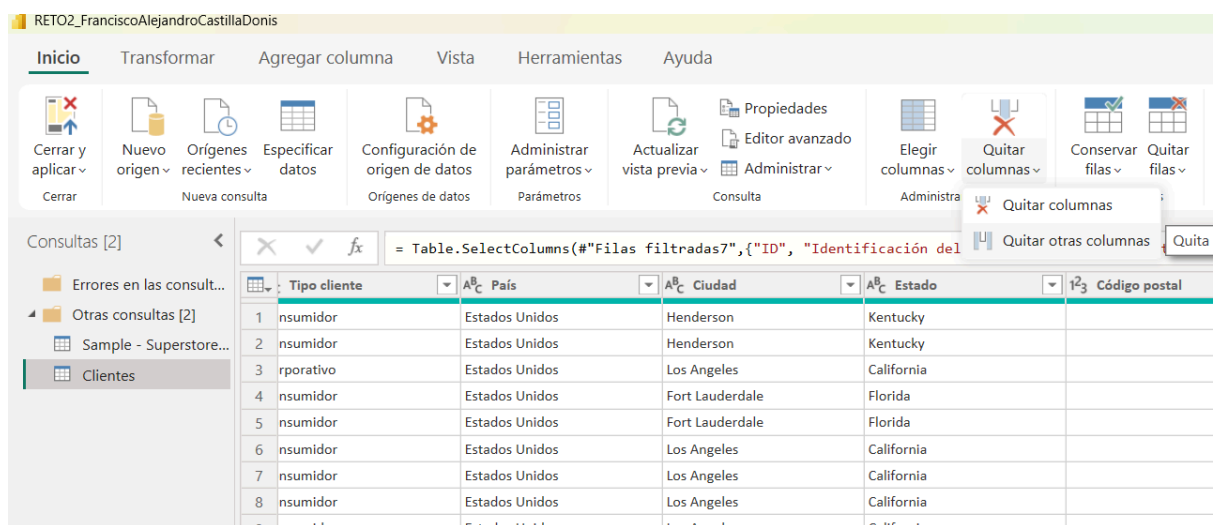
4

Se han creado nuevas tablas, dividiendo las columnas por sus características y datos aportados



Seleccionamos las columnas que queremos y quitamos el resto columnas

Inicio → Quitar columnas → quitar otras columnas y llamamos a esta nueva tabla “clientes”



Duplicamos → y creamos otra tabla (llamada Productos) que contiene: ID producto, nombre producto, categoria, subcategoria

Vamos a crear la tabla calendario:

Salimos de Query y nos vamos al modelado

Herramientas tabla → nueva tabla → la creamos con DAX

Archivo Inicio Ayuda **Herramientas de tablas**

Nombre Sample - Supersto...

Administrar relaciones Relaciones

Nueva medida Nueva medida rápida Cálculos

Nueva columna Nueva tabla

Marcar como 1 fechas Calendario

Estructura

ID	Orden ID	Fecha de pedido	Fecha de envío	Modo de embarque	Ider
43	CA-2016-101343	7/17/2016	7/22/2016	Clase estándar	RA-1
514	CA-2017-163405	12/21/2017	12/25/2017	Clase estándar	BN-
515	CA-2017-163405	12/21/2017	12/25/2017	Clase estándar	BN-
1606	US-2016-115819	4/19/2016	4/24/2016	Segunda Clase	JO-1
1607	US-2016-115819	4/19/2016	4/24/2016	Segunda Clase	JO-1
1609	US-2016-115819	4/19/2016	4/24/2016	Segunda Clase	JO-1
1807	US-2016-116729	12/25/2016	12/28/2016	Primera Clase	GK-1
2060	CA-2014-106439	10/31/2014	11/4/2014	Clase estándar	GG-

Dim_Calendario =

ADDCOLUMNS(

CALENDAR(

DATE(2020, 1, 1), // Fecha de inicio (ajusta según tus datos)

DATE(2025, 12, 31) // Fecha de fin (ajusta según necesites)

),

"Año", YEAR([Date]),

"Número de Año", YEAR([Date]),

"Trimestre", "Q" & QUARTER([Date]),

"Número de Trimestre", QUARTER([Date]),

"Mes", FORMAT([Date], "MMMM"),

"Nombre Mes Corto", FORMAT([Date], "MMM"),

"Número de Mes", MONTH([Date]),

"Semana del Año", WEEKNUM([Date], 2),

"Día del Mes", DAY([Date]),

"Día de la Semana", FORMAT([Date], "dddd"),

"Día de Semana Corto", FORMAT([Date], "ddd"),

"Número Día Semana", WEEKDAY([Date], 2),

"Día del Año", DATEDIFF(DATE(YEAR([Date]), 1, 1), [Date], DAY) + 1,

"Mes-Año", FORMAT([Date], "MM/yyyy"),

"Trimestre-Año", "Q" & QUARTER([Date]) & "-" & YEAR([Date]),

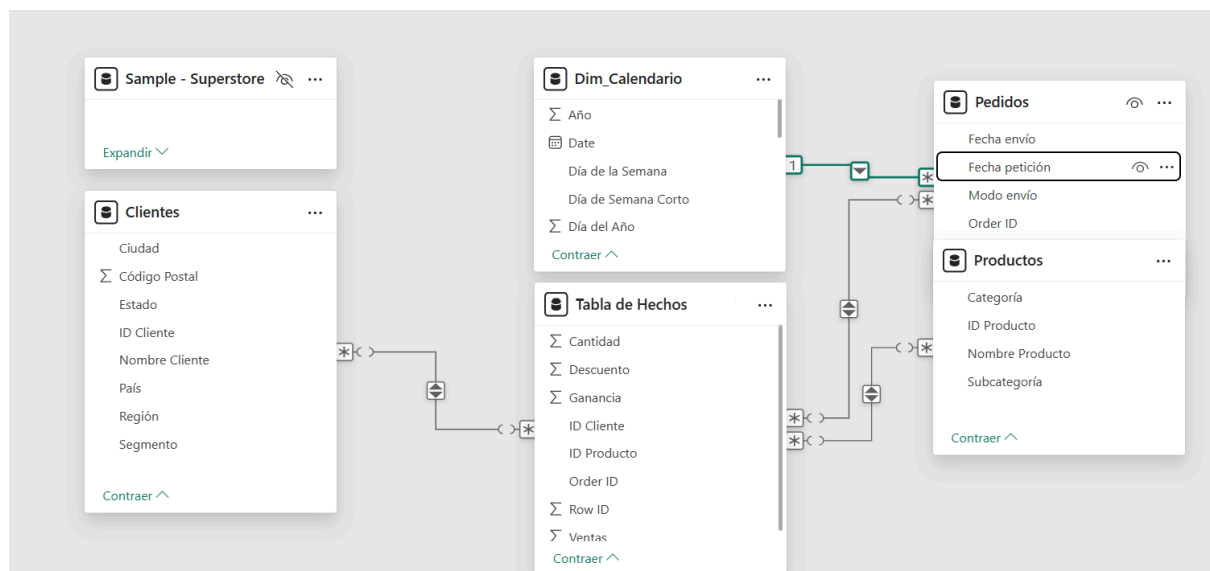
"Es Fin de Semana", IF(WEEKDAY([Date], 2) > 5, "Sí", "No"),

"Día Laborable", IF(WEEKDAY([Date], 2) < 6, "Sí", "No"),

"FechaKey", YEAR([Date]) * 10000 + MONTH([Date]) * 100 + DAY([Date])

)

Crear relaciones entre las nuevas tablas



Pasamos a **CREAR MEDIDAS**

Medidas en la tabla de hechos:

- Ventas totales → Ventas totales = `SUM ('Tabla de Hechos'[Ventas]`
- Beneficio Total → Beneficio Total = `SUM ('Tabla de Hechos'[Ganancia]`
- Margen beneficio → Margen Beneficio = `DIVIDE([Beneficio Total], [Ventas totales])`
- Cantidad Total Vendida → Cantidad Total Vendida = `SUM('Tabla de Hechos'[Cantidad])`
- Descuento promedio → Descuento Promedio = `AVERAGE('Tabla de Hechos'[Descuento])`
-

VISUALIZACIONES

- Incorporar un **gráfico de tabla**
detallamos un listado de productos y las ventas de cada uno (pudiendo realizar una suma de esa columna de ventas)
- Incluimos una **tarjeta**
- **Gráfico de línea**
incluyendo en el eje dos campos, ej: año y trimestre
- **Mapa**
Archivo → Configuración → global → Seguridad se puede cambiar el tamaño de las marcas del mapa
- **Slicer** para realizar filtros
- Insertamos títulos con **cuadros de texto**

Adaptamos el color de fondo acorde al branding de la empresa

Crear un logotipo para la empresa y lo añadimos (*Insertar → imagen*)

