



700 Fruve Access. Presentación. Fase I: Proceso ETL y Modelo de Datos

Diseño e implantación de un proyecto de Análisis de Datos y de Inteligencia de Negocio para FRUVE

Jose Ignacio González Gómez

Departamento de Economía, Contabilidad y Finanzas - Universidad de La Laguna

www.jggomez.eu

V.2.4

Ejercicio Basado: [David Asurmendi](#)

Proyecto de analítica de datos

Coordinador técnico: J. Ignacio Glez

Contenido

1	Presentación del proyecto e información disponible	2
1.1	Objetivo, fases y requisitos del proyecto	2
1.2	Información disponible	2
2	Fase I. Se pide	3
2.1	Proceso ETL	3
2.2	Creacion de modelo de datos	3
2.2.1	Crear tabla calendario	3
2.2.2	Relaciones entre las tablas	4
2.2.3	Crear tabla para organizar las medidas	4

1 Presentación del proyecto e información disponible

1.1 Objetivo, fases y requisitos del proyecto

Hemos aceptado el desarrollo de un proyecto de analítica de datos para la empresa FRUVE cooperativa agrícola dedicada a la comercialización de productos perecederos de frutas y verduras que opera en diversos países europeos dedicada a la comercialización mayorista de productos perecederos.

Este encargo consta de un presupuesto cerrado que tiene las siguientes fases:

1. Fase I Desarrollo de las bases del modelo. Proceso ETL y el diseño del modelo de datos
2. Fase II Analítica Básica I: Productos, Clientes y Comerciales. Funciones Básicas
3. Fase III Analítica Básica II: Productos, Clientes y Comerciales. Funciones de Filtro
4. Fase IV Analítica de Tiempos. Cumplimiento de pedidos y análisis temporal. Funciones de fecha y hora y funciones de inteligencia de tiempo

Todas las propuestas relacionada con la analítica de datos, tablas y graficos deben estar basados en Excel por requerimiento expreso de la dirección de la empresa, por tanto nuestros diseños se basaran en el uso de tablas y graficos dinámicos, segmentadores, etc.. fundamentados todos ellos en un modelo de datos como el propuesto inicialmente y que podrá ampliarse a medida que la demanda y requerimientos de la empresa asi lo exija.

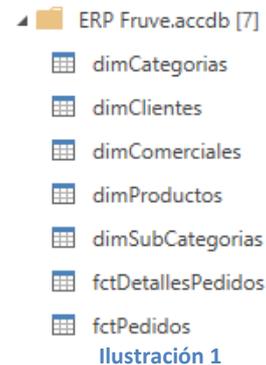
1.2 Información disponible

Contamos con una base de datos en ACCESS que contiene la información relacionada con el ERP de la empresa con siete tablas que deben examinarse previo a crear la conexión y posterior incorporación al modelo de datos.

Señalar que disponemos de dos tablas de hechos, la de pedidos (fctPedidos) y la de detalle de pedidos (fctDetallesPedidos) siendo el resto tablas de dimensión.

Como primera aproximación, independientemente del analisis pormenorizado de cata tabla el contenido de cada una de ellas es el siguiente:

1. dimCategorías, tabla simple que contiene las categorías de los productos.
2. dimClientes. Contiene información relacionada con el propio nombre de la tabla.
3. dimComerciales, información relacionada.
4. fctDetallePedidos, contiene el detalle de cada uno de los pedidos (producto, cantidad, precio, descuento, etc)
5. fctPedidos, información básica del pedido como es la fecha, cliente, comercial...
6. dimProductos, información relacionada.
7. dimSubcategorías, información relacionada.



2 Fase I. Se pide

Elaborar los siguientes informes y cuadros de mando basado en un modelo de datos con medidas basicas y de inteligencia de tiempo.

2.1 Proceso ETL

En este caso proceso ETL es sencillo ya que toda la información se encuentra correctamente estructurada en una base de datos como hemos expuesto anteriormente y por tanto esta correctamente estructurada en tablas.

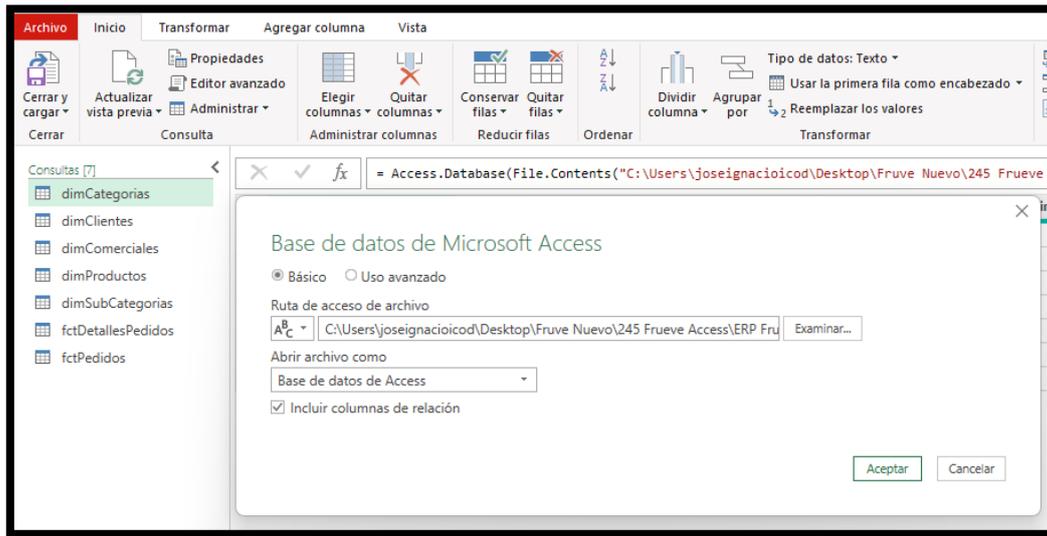


Ilustración 2

La limpieza de los datos así como la estructura hace que no sea necesario llevar a cabo ningún proceso de transformación por tanto estos serán cargados directamente al modelo.

2.2 Creacion de modelo de datos

2.2.1 Crear tabla calendario

Ver: [Anexo Calendario](#)

Las fechas juegan un papel estratégico no solo para el análisis de la perspectiva temporal sobre la evolución de determinadas variables como el número de pedidos y mercados sino especialmente en el análisis y control estratégico del tiempo de cumplir con los pedidos (diferencia fechas de inicio y cierre del pedido).

Por ello proponemos crear una tabla calendario similar a la siguientes con las siguientes columnas calculadas (en negrita).

	Date	Año	NMes	Mes	MMM-YYYY	NDiaSe...	DiaSemana	Trimestre	Quincena	NSemana	Bimestre
1	31/12/2020 0:00:00	2020	12	diciembre	dic-2020		5 jueves	4-T	2ª Quincena		53 6-Bimestre
2	30/12/2020 0:00:00	2020	12	diciembre	dic-2020		4 miércoles	4-T	2ª Quincena		53 6-Bimestre
3	29/12/2020 0:00:00	2020	12	diciembre	dic-2020		3 martes	4-T	2ª Quincena		53 6-Bimestre
4	28/12/2020 0:00:00	2020	12	diciembre	dic-2020		2 lunes	4-T	2ª Quincena		53 6-Bimestre

Ilustración 3

Debemos garantizar que el rango de fechas debe estar entre el valor mínimo (01/01/2022) y el valor máximo (31/12/2024) del valor de fecha de inicio y cierre de los pedidos.

2.2.2 Relaciones entre las tablas

Una vez agregada las tablas basicas al modelo queda establecer las relaciones tal y como proponemos a continuación.

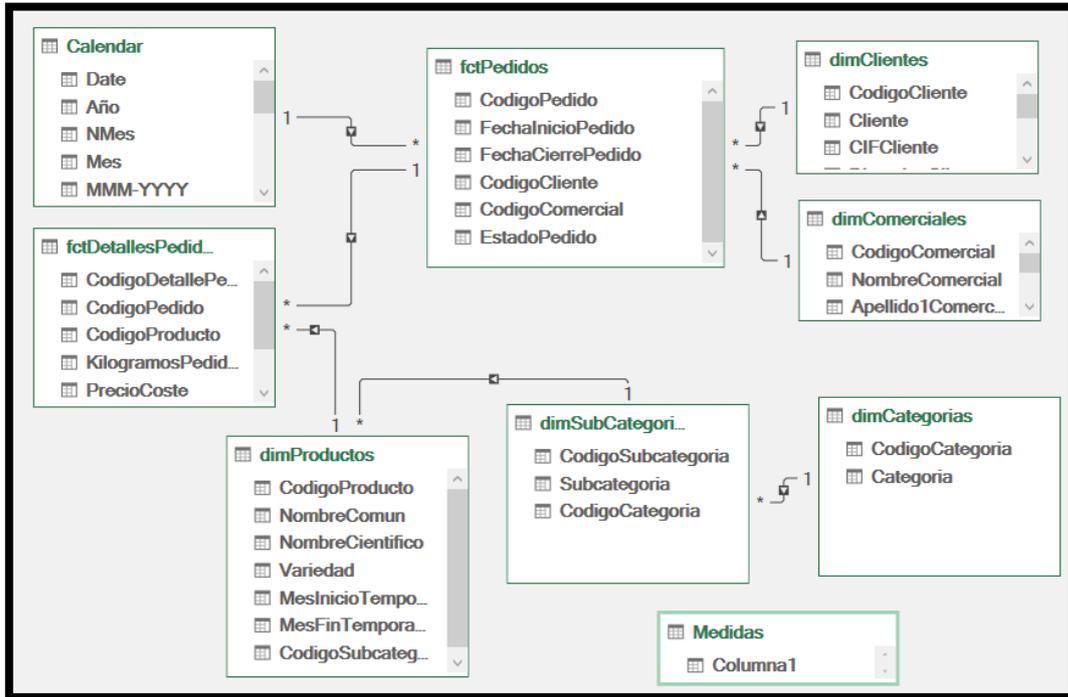


Ilustración 4

Administrar relaciones

Crear Editar Eliminar

Activo	Tabla 1	Cardinalidad	Dirección de filtro	Tabla 2
Si	dimProductos [CodigoSubcategoria]	Varios a uno (*:1)	<< A dimProductos	dimSubCategorias [CodigoSubcategoria]
Si	dimSubCategorias [CodigoCategoria]	Varios a uno (*:1)	<< A dimSubCategorias	dimCategorias [CodigoCategoria]
Si	fctDetallesPedidos [CodigoPedido]	Varios a uno (*:1)	<< A fctDetallesPedidos	fctPedidos [CodigoPedido]
Si	fctDetallesPedidos [CodigoProducto]	Varios a uno (*:1)	<< A fctDetallesPedidos	dimProductos [CodigoProducto]
Si	fctPedidos [CodigoCliente]	Varios a uno (*:1)	<< A fctPedidos	dimClientes [CodigoCliente]
Si	fctPedidos [CodigoComercial]	Varios a uno (*:1)	<< A fctPedidos	dimComerciales [CodigoComercial]
Si	fctPedidos [FechaInicioPedido]	Varios a uno (*:1)	<< A fctPedidos	Calendar [Date]

Ilustración 5

2.2.3 Crear tabla para organizar las medidas

Con el fin de organizar las medidas crearemos en Excel una tabla en blanco que añadiremos al modelo.