

BASES DE DATOS II

POR:

SEBASTIAN BEDOYA GIRALDO

GRUPO:

PREICA2401B010093

ACTIVIDAD 1

EA1: MODELO ESTRELLA DE UN DATA MART

PROFESOR:

VICTOR HUGO MERCADO RAMOS

INSTITUCIÓN UNIVERSITARIA DIGITAL DE ANTIOQUIA

2024

Introducción

En la situación actual, es crucial para las empresas tener una gestión eficiente de datos para lograr el éxito. Poder analizar y comprender la información de manera efectiva ayuda a tomar decisiones estratégicas informadas y optimizar las operaciones. En este sentido, el modelo estrella se presenta como una herramienta esencial para la empresa "Jardinería", siendo ampliamente utilizada en el ámbito de la inteligencia empresarial (BI) como estructura base de datos.

Se enfoca en el análisis de ventas la presentación de este documento, que muestra el diseño e implementación de un modelo estrella para la base de datos "Jardinería". El modelo estrella seleccionado facilita la exploración de las relaciones entre variables como productos, clientes, tiempo y empleados para identificar tendencias, patrones de compra y áreas de oportunidad.

Objetivos

1. Desarrollar un modelo estrella eficiente: Construir un modelo estrella que permita a la empresa "Jardinería" analizar y comprender eficientemente sus datos de ventas es el objetivo principal.
2. Identificar información clave: Se desea identificar y analizar tres categorías esenciales para la empresa.
 - Producto estrella: Encontrar el producto más demandado, teniendo en cuenta las unidades vendidas.
 - Categoría con mayor diversidad: Hallar cuál es la categoría que tiene la mayor variedad de productos.
 - Año de mayor éxito: Determinar el año en que se registraron las ventas más altas.
3. Optimizar la toma de decisiones: Basarse en información precisa y análisis multidimensionales para proporcionar una base sólida de toma de decisiones estratégicas en el futuro.

Planteamiento del Problema

La base de datos actual de "Jardinería" está en un formato normalizado, lo que hace difícil el análisis directo y la extracción de información valiosa a pesar de asegurar la integridad de los datos. Este tipo de estructura provoca ineficiencia en las consultas complejas y multidimensionales. Es necesario transformar la base de datos a un modelo estrella para facilitar el análisis de las ventas y obtener respuestas a las tres categorías de interés.

Análisis del Problema

Las diferentes tablas tienen relaciones de muchos a uno en la estructura actual de la base de datos, con "Pedido" como el núcleo central. Se analizarán las tablas y sus relaciones para identificar los campos esenciales necesarios para la creación del modelo estrella, que incluye la

tabla de hechos y las dimensiones. La información de ventas, clientes, productos, fechas, empleados, oficinas y categorías de productos será considerada.

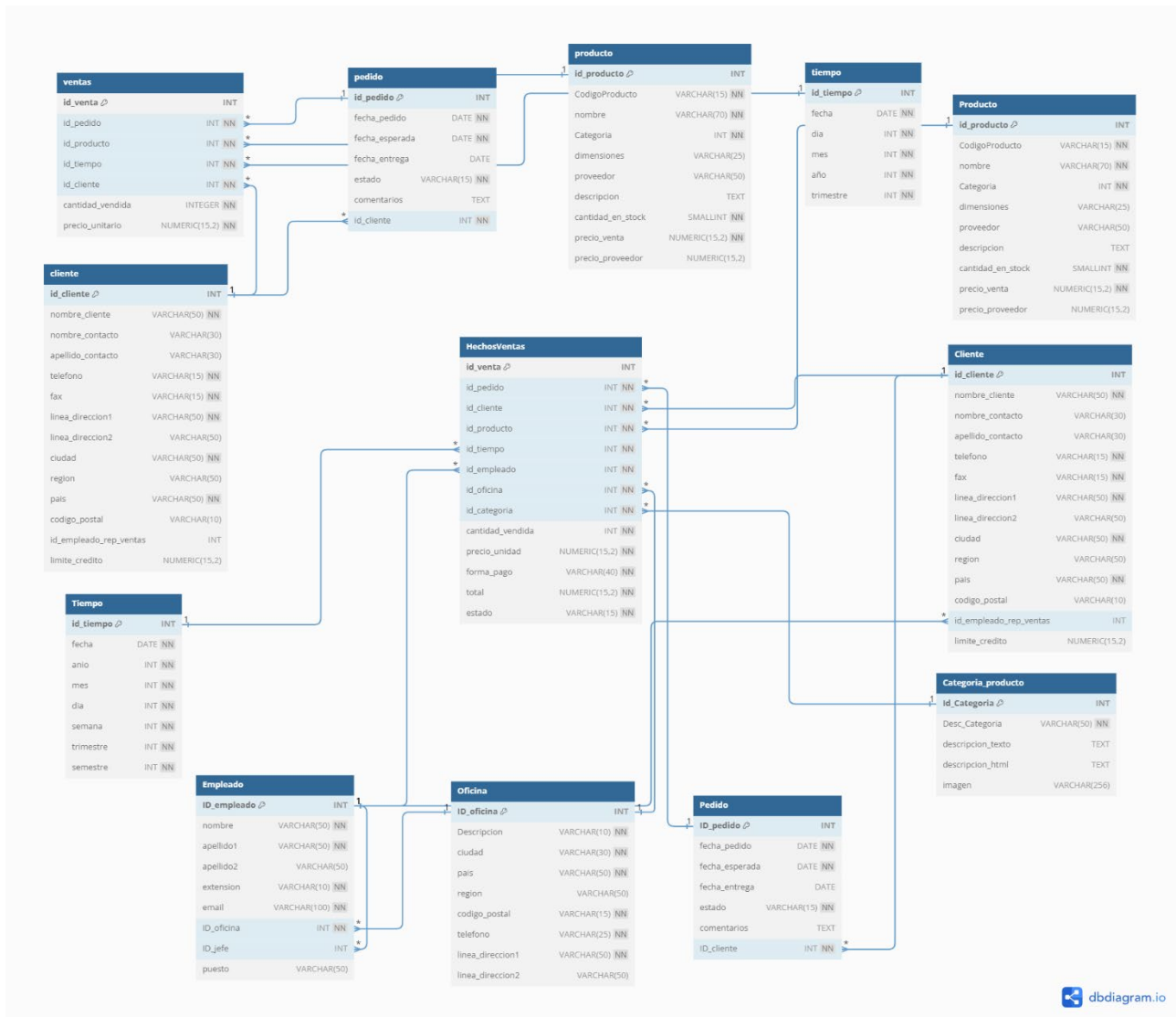
Propuesta de la Solución: Modelo Estrella "Jardinería"

1. Descripción del Modelo

El modelo estrella propuesto para "Jardinería" se compone de una tabla de hechos central denominada "HechosVentas" y seis dimensiones que proporcionan contexto a los hechos: Cliente, Producto, Tiempo, Empleado, Oficina y Categoría_producto.

- **Tabla de Hechos (HechosVentas):** Esta tabla contiene información detallada sobre las transacciones de venta, incluyendo:
 - id_venta: Clave primaria que identifica de manera única cada venta.
 - id_pedido: Clave foránea que relaciona la venta con el pedido original.
 - id_cliente: Clave foránea que conecta la venta con el cliente que realizó la compra.
 - id_producto: Clave foránea que vincula la venta con el producto vendido.
 - id_tiempo: Clave foránea que relaciona la venta con la fecha en que se realizó.
 - id_empleado: Clave foránea que asocia la venta con el empleado que la gestionó.
 - id_oficina: Clave foránea que conecta la venta con la oficina desde la que se realizó.
 - id_categoria: Clave foránea que relaciona la venta con la categoría del producto vendido.
 - cantidad_vendida: Cantidad de unidades del producto vendidas en la transacción.
 - precio_unidad: Precio unitario del producto en el momento de la venta.
 - forma_pago: Método de pago utilizado por el cliente.
 - total: Importe total de la venta.
 - estado: Estado actual del pedido (entregado, pendiente, rechazado, etc.).
- **Dimensiones:**
 - Cliente: Almacena información detallada de los clientes, como nombre, datos de contacto, dirección, límite de crédito, etc.
 - Producto: Contiene información sobre los productos, incluyendo código, nombre, categoría, dimensiones, proveedor, descripción, cantidad en stock, precio de venta y precio del proveedor.
 - Tiempo: Permite analizar las ventas en función del tiempo, con información como fecha, día, mes, año, trimestre y semestre.
 - Empleado: Almacena información sobre los empleados que gestionan las ventas, incluyendo nombre, apellidos, extensión, correo electrónico, oficina y puesto.
 - Oficina: Contiene información sobre las oficinas, como la descripción, ciudad, país, región y código postal.
 - Categoría_producto: Clasifica los productos en diferentes categorías y proporciona información adicional como descripción.

2. Diseño del Modelo Estrella



3. Implementación

Para implementar el modelo estrella en MySQL, se seguirán los siguientes pasos:

1. Creación de las tablas: Se ejecutará el script SQL proporcionado anteriormente para crear la tabla de hechos "HechosVentas" y la dimensión "Tiempo", así como las dimensiones existentes en la base de datos "Jardinería".
2. Población de las tablas: Se utilizarán herramientas ETL (extracción, transformación y carga) para migrar los datos de la base de datos original al nuevo modelo estrella. Esto incluye:
 - Extraer la información relevante de las tablas "Pedido", "detalle_pedido" y "pago" y transformarla para integrarla en la tabla de hechos "HechosVentas".
 - Generar la dimensión "Tiempo" con las fechas y los desgloses de tiempo necesarios para el análisis.
 - Poblar las demás dimensiones con la información de las tablas correspondientes.

3. Verificación del modelo: Se verificará que las tablas estén creadas correctamente y que las relaciones entre la tabla de hechos y las dimensiones sean las adecuadas.

4. Análisis y Visualización.

Una vez que el modelo estrella esté implementado y poblado con datos, se podrán realizar consultas multidimensionales para analizar las ventas desde diferentes perspectivas.

1. Producto estrella: Se pueden realizar consultas para identificar el producto con mayor cantidad vendida, lo que indicaría el producto con mayor demanda.
2. Categoría con mayor diversidad: Se pueden realizar consultas para contar la cantidad de productos distintos en cada categoría y así determinar la categoría con mayor variedad.
3. Año de mayor éxito: Se pueden agrupar las ventas por año y calcular el total de ventas para cada año, identificando así el año con mayor volumen de ventas.

Además, se pueden utilizar herramientas de visualización de datos para crear gráficos y dashboards que permitan visualizar las tendencias y patrones de compra de manera clara e intuitiva.

Conclusiones:

El modelo estrella implementado para la base de datos "Jardinería" proporciona una solución eficaz y adaptable para el análisis de ventas. La empresa puede identificar información clave sobre sus productos, clientes y mercado mediante este modelo, lo que hace más sencilla la toma de decisiones estratégicas informadas y la optimización de las operaciones.

El modelo estrella es una herramienta fundamental para mejorar la gestión de datos en "Jardinería" e impulsar su crecimiento en el mercado.

Anexos:

<https://dbdiagram.io/d/6625d53d03593b6b6194902a>

Bibliografía:

Kimball, R., & Ross, M. (2013). The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling (3rd ed.). Wiley.

Sánchez, M. G. (2023, 19 julio). Beneficios del Modelo Estrella en las empresas.
<https://www.tecon.es/que-es-el-modelo-estrella/>

BASES DE DATOS II

POR:

SEBASTIAN BEDOYA GIRALDO

GRUPO:

PREICA2401B010093

ACTIVIDAD 2

EA2: CREACIÓN DE UNA BASE DE DATOS DE STAGING

PROFESOR:

VICTOR HUGO MERCADO RAMOS

INSTITUCIÓN UNIVERSITARIA DIGITAL DE ANTIOQUIA

2024

Introducción

En el mundo empresarial competitivo de hoy, el análisis de datos es fundamental para alcanzar el éxito. Transformar datos sin procesar en información procesable es fundamental para impulsar la toma de decisiones estratégicas, optimizar operaciones y lograr un mayor rendimiento.

Dado que es crucial, reconocemos la importancia de gestionar los datos eficientemente y por ello hemos iniciado el desarrollo de una base de datos para Jardinería. El propósito de esta iniciativa es mejorar la calidad de los datos, optimizar su estructura para el análisis y facilitar la generación oportuna y precisa de informes.

Objetivos

Nuestros objetivos al construir la base de datos de staging son: Nuestros objetivos al construir la base de datos de staging son:

- Mejorar la calidad de los datos: Garantizar la precisión de los datos al eliminar redundancias, inconsistencias y valores faltantes.
- Optimizar la estructura de datos: Mejorar el rendimiento del análisis y la generación de informes mediante la implementación de un modelo de estrella.
- Facilitar el análisis de datos: Posibilitar la extracción eficaz y veloz de información relevante de los datos.
- Mejorar la toma de decisiones: Brindar información exacta y puntual para apoyar la toma de decisiones estratégicas y operativas.

Planteamiento del problema

La analítica de datos es imprescindible para el éxito de cualquier empresa en la actualidad. No obstante, la estructura y organización de las bases de datos existentes a menudo dificultan la capacidad para extraer información significativa de los datos.

La base de datos Jardinería, no es adecuada para un análisis eficiente de datos ya que está diseñada principalmente para operaciones transaccionales. Generar informes y obtener insights accionables se dificulta debido a su estructura rígida, la presencia de redundancias e inconsistencias, y la falta de un modelo de datos orientado al análisis.

Necesidades:

Para superar estos desafíos, se requiere la creación de una base de datos de staging que:

Implemente un modelo de datos orientado al análisis: Utilizar un modelo de estrella para hacer más fácil la agregación de datos, el análisis multidimensional y la generación eficiente de informes

Elimine redundancias e inconsistencias: Asegurar la calidad e integridad de la información mediante la implementación de procesos de limpieza y transformación de datos

Ofrezca flexibilidad para el análisis ad-hoc: Facilitar la generación de ideas inesperadas y respuestas a preguntas emergentes del negocio permitiendo explorar los datos desde varias perspectivas.

Crear una base de datos de staging para Jardinería no consiste en copiar la información actual, sino en modificarla y organizarla para que se convierta en un recurso estratégico para tomar decisiones.

Análisis del problema

Realizamos un análisis en profundidad de la base de datos Jardinería para comprender su estructura, identificar áreas de mejora y determinar los requisitos para la base de datos de staging. Este análisis reveló lo siguiente:

Tablas principales: Las tablas "Producto", "Cliente", "Tiempo", "Pedido" y "Empleado" son esenciales para el análisis, ya que contienen información sobre las transacciones, los clientes, los empleados y los productos, respectivamente.

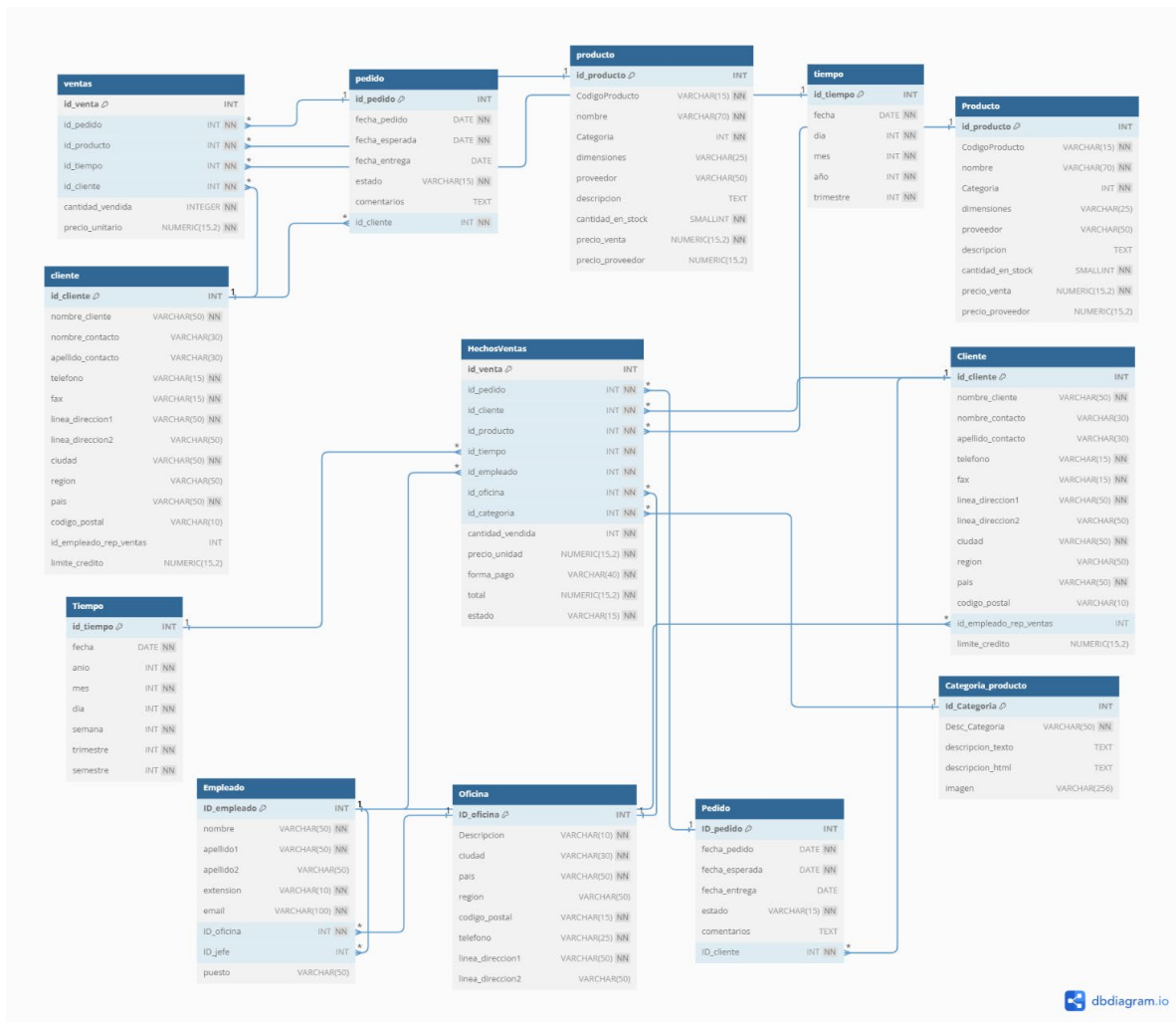
Relaciones entre tablas: Comprender las relaciones entre las tablas, como la relación uno a muchos entre "Pedido" y "Cliente", es crucial para definir las claves externas en la base de datos de staging.

Redundancias e inconsistencias: Se encontraron instancias de datos duplicados e inconsistencias en varias tablas, lo que destaca la necesidad de limpieza y transformación de datos durante el proceso de staging.

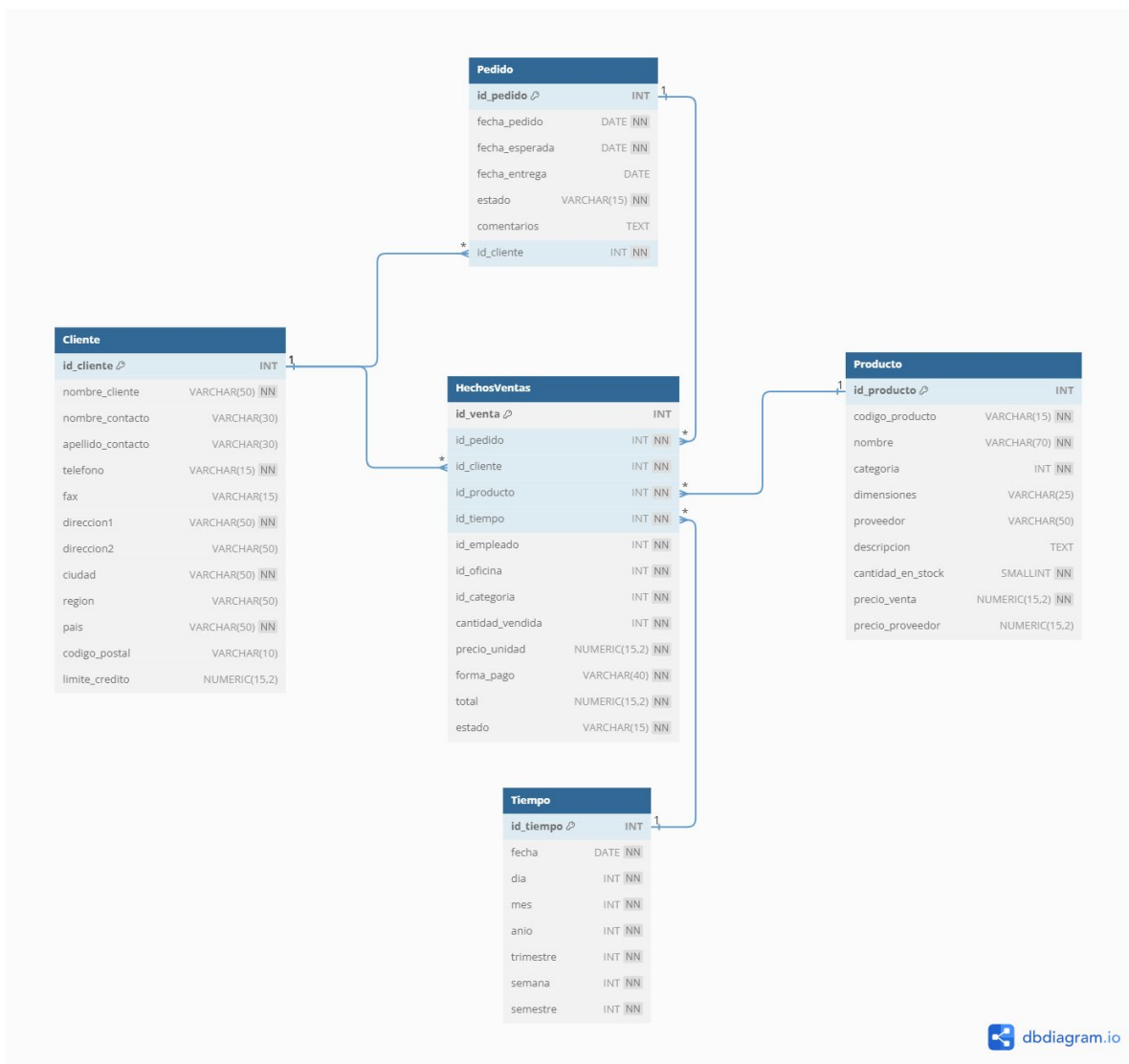
Propuesta de solución

Modelo de estrella:

MODELO ESTRELLA – PRIMERA ENTREGA



MODELO ESTRELLA – JARDINERIA CORREGIDA



Se elige un modelo estrella para la base de datos de staging ya que proporciona una estructura optimizada para analizar los datos. La tabla "**HECHOSVENTAS**" es el núcleo de este modelo y contiene las métricas clave del negocio, tales como la cantidad de pedidos, el precio total y el descuento. Esta tabla de hechos está conectada a varias tablas de dimensiones, incluyendo:

Dim_Cliente: Muestra la fecha del pedido con sus atributos de año, mes y día.

Dim_producto: Incluye datos de los clientes, como nombre, dirección y detalles para contactar.

Dim_tiempo: Guarda información de los empleados, incluyendo su nombre, puesto y fecha de inicio.

Dim_pedido: Proporciona información sobre los productos, incluyendo el nombre, categoría, precio y cantidad disponible.

Dim_HechosVentas: Captura las transacciones de ventas, actuando como la tabla de hechos central en el modelo de datos estrella.

Descripción del análisis realizado a los datos Jardinería y cómo estos se trasladaron a la base de datos Staging

Se llevó a cabo un análisis exhaustivo de cada tabla en la base de datos Jardinería para identificar los campos relevantes para el modelo de estrella.

Antes de cargar los datos en la base de datos de preparación, se realizaron consultas SQL para extraer, limpiar y transformarlos.

Ejemplos de Consultas SQL:

Producto

Campos: id_producto, nombre, categoria, dimensiones, proveedor, descripcion, cantidad_en_stock, precio_venta, precio_proveedor.

```
SELECT
  id_producto,
  nombre,
  categoria,
  dimensiones,
  proveedor,
  descripcion,
  cantidad_en_stock,
  precio_venta,
  precio_proveedor
FROM
  Producto
ORDER BY
  id_producto ASC;
```

Cliente

Campos: id_cliente, nombre_cliente, nombre_contacto, apellido_contacto, telefono, fax, ciudad, region, pais, codigo_postal, limite_credito.

```

SELECT
  id_cliente,
  nombre_cliente,
  nombre_contacto,
  apellido_contacto,
  telefono,
  fax,
  ciudad,
  region,
  pais,
  codigo_postal,
  limite_credito
FROM
  Cliente
ORDER BY
  id_cliente ASC;

```

Tiempo

Campos: id_tiempo, fecha, día, mes, anio, trimestre, semana, semestre.

```

| SELECT
  id_tiempo,
  fecha,
  dia,
  mes,
  anio,
  trimestre,
  semana,
  semestre
FROM
  Tiempo
ORDER BY
  fecha ASC;

```

Pedido

Campos: id_pedido, fecha_pedido, fecha_esperada, fecha_entrega, estado, comentarios, id_cliente. También se incluyen detalles de Cliente para relacionar cada pedido con la información del cliente.

```

SELECT
  p.id_pedido,
  p.fecha_pedido,
  p.fecha_esperada,
  p.fecha_entrega,
  p.estado,
  p.comentarios,
  p.id_cliente,
  c.nombre_cliente
FROM
  Pedido p
JOIN
  Cliente c ON p.id_cliente = c.id_cliente
ORDER BY
  p.id_pedido ASC;

```

Además, también se utilizó una tabla de hechos en tus consultas:

HechosVentas

Campos: id_venta, cantidad_vendida, precio_unidad, forma_pago, total, estado, además de enlaces a las dimensiones Producto, Cliente, Tiempo, Empleado y Oficina.

```

SELECT
  hv.id_venta,
  hv.cantidad_vendida,
  hv.precio_unidad,
  hv.forma_pago,
  hv.total,
  hv.estado,
  p.nombre AS producto_nombre,
  c.nombre_cliente,
  t.fecha,
  e.nombre AS empleado_nombre,
  o.descripcion AS oficina_descripcion
FROM
  HechosVentas hv
JOIN
  Producto p ON hv.id_producto = p.id_producto
JOIN
  Cliente c ON hv.id_cliente = c.id_cliente
JOIN
  Tiempo t ON hv.id_tiempo = t.id_tiempo
JOIN
  empleado e ON hv.id_empleado = e.id_empleado
JOIN
  oficina o ON hv.id_oficina = o.id_oficina

```

- Anexos: debe agregar los BK de las dos bases de datos y el documento del script de las consultas para crear la base de datos Staging.

Object Explorer

Connect

- LocalDb\jardineria (SQL Server 15.0.41)
- System Databases
- Database Snapshots
- Datos_Staging
- Database Diagrams
- Tables
- System Tables
- FileTables
- External Tables
- Graph Tables
- dbo.Categoria_producto
- dbo_cliente
- dbo_detalle_pedido
- dbo_empleado
- dbo.HechosVentas
- dbo_oficina
- dbo_pago
- dbo_pedido
- dbo_producto
- dbo_Tiempo
- Views
- External Resources
- Synonyms
- Programmability
- Service Broker
- Storage
- Security
- jardineria
- Security
- Server Objects
- Replication
- PolyBase
- Management
- XE-Event Profiler

SQL Query1.sql - (L:\45CR6T\SEBAS (70))*

SQL Query1.sql - (L:\45CR6T\SEBAS (80))*

100%

Results **Messages**

id_producto	nombre	categoria	dimensiones	proveedor	descripcion	cantidad_en_stock	precio_venta	precio_proveedor
1	Sierra de Poda 400MM	2	0.258	HiperGarden Tools	Gracias a la poda se consigue manipular un poco ...	15	14.00	11.00

id_cliente	nombre_cliente	nombre_contacto	apellido_contacto	telefono	fax	ciudad	region	pais	codigo_postal	limite_credito
1	GoldFish Garden	Daniel G	GoldFish	5556901745	5556901746	San Francisco	NULL	USA	24006	3000.00
2	Gardening Associates	Anee Wright	Wright	5557410345	5557410346	Miami	NULL	USA	24010	6000.00
3	Garden Valley	Lark	Rhale	5552201129	5552201128	New York	NULL	USA	25405	12000.00
4	Tendo Garden	Alane	Tendo	55591233210	55591233211	Miami	NULL	USA	696569	60000.00
5	Lasas S.A	Antonio	Lasas	34916540145	34914851312	Fuenlabrada	Madrid	Spain	28945	154310.00
6	Beragua	Jose	Bempejo	645487321	916549872	Madrid	Madrid	Spain	28942	20000.00
7	Club Golf Puerta del hierro	Paco	Lopez	62456810	919535678	Madrid	Madrid	Spain	28930	40000.00
8	Naturagua	Guillermo	Rengifo	689234750	916428956	Madrid	Madrid	Spain	28947	32000.00

id_tiempo	fecha	dia	mes	año	trimestre	semana	semestre
1	2021-02-15	15	1	2021	1	3	
2	2021-04-10	10	4	2021	2	15	1
3	2021-07-05	5	7	2021	3	27	2
4	2021-10-25	25	10	2021	4	43	2

id_pedido	fecha_pedido	fecha_esperada	fecha_entrega	estado	comentarios	id_cliente	nombre_cliente
16	2009-01-18	2009-02-12	NULL	Pendiente	entregar en murcia	9	DavaDistribuc...
17	2009-01-20	2009-02-15	NULL	Pendiente	NULL	9	DavaDistribuc...
18	2009-01-09	2009-01-09	2009-01-09	Rechaz.	mal pago	9	DavaDistribuc...
19	2009-01-11	2009-01-11	2009-01-11	Entregado	NULL	9	DavaDistribuc...
20	2008-12-30	2009-01-10	NULL	Rechaz.	El pedido fue anulado por el cliente	5	Lasas S.A
21	2008-07-14	2008-07-31	2008-07-25	Entregado	NULL	14	Jardín de Flor...
22	2009-02-02	2009-02-08	NULL	Rechaz.	El cliente carece de saldo en la cue...	1	GoldFish Gar...
23	2009-02-06	2009-02-12	NULL	Rechaz.	El cliente anula la operacion para a...	3	Gerardo Valley

id_venta	cantidad_venta	precio_unidad	foma_pago	total	estado	producto_nombre	nombre_cliente	fecha	empleado_nombre	oficina_descripcion
----------	----------------	---------------	-----------	-------	--------	-----------------	----------------	-------	-----------------	---------------------

Anexos:

<https://dbdiagram.io/d/6625d53d03593b6b6194902a>

Conclusión

La capacidad de la empresa para analizar datos y obtener información estratégica se mejora significativamente con la construcción de una base de datos de staging para Jardinería. Durante el proceso de staging, la implementación del modelo estrella, la limpieza de datos y su transformación sientan las bases sólidas para una toma de decisiones informada, optimización operativa y crecimiento empresarial.

Bibliografía

Kimball, R., & Ross, M. (2013). The data warehouse toolkit: The definitive guide to dimensional modeling (3rd ed.). John Wiley & Sons. Este libro es una referencia fundamental para el modelado dimensional, que se utiliza a menudo en las bases de datos de staging.

Loshin, D. (2011). *Business intelligence: The savvy manager's guide* (2nd ed.). Morgan Kaufmann. Cubre los conceptos básicos de Business Intelligence, incluyendo el papel de las bases de datos de staging.

Han, J., Pei, J., & Kamber, M. (2011). Data mining: Concepts and techniques (3rd ed.). Morgan Kaufmann.

BASES DE DATOS II

POR:

SEBASTIAN BEDOYA GIRALDO

GRUPO:

PREICA2401B010093

ACTIVIDAD 3

EA3: PROCESO DE TRANSFORMACIÓN DE DATOS Y CARGA EN EL DATA MART FINAL

PROFESOR:

VICTOR HUGO MERCADO RAMOS

INSTITUCIÓN UNIVERSITARIA DIGITAL DE ANTIOQUIA

2024

Introducción

En la actualidad, el manejo de información se ha vuelto fundamental para impulsar decisiones estratégicas y alcanzar un funcionamiento óptimo. En este contexto, resulta indispensable la creación de una base de datos staging con el fin de asegurar que los datos estén completos, sean confiables y estén disponibles para su posterior análisis.

El proceso empieza al revisar el modelo estrella que se ha definido para el proyecto de análisis de la base de datos, donde se incluyen tanto la estructura como las relaciones entre las tablas dimensionales y la tabla factual. El Data mart final utiliza la base de datos de staging, que se ha creado previamente, como una importante área temporal para transformar y limpiar los datos extraídos antes de ser cargados.

Se proporcionan a continuación las consultas SQL necesarias para extraer los datos, asegurando su integridad y coherencia. Después, se implementan métodos de transformación que incluyen la limpieza y normalización de los datos. Por último, se procede a ejecutar las consultas de carga para insertar los registros transformados en el Data mart final y se verifica que los datos hayan sido correctamente ingresados.

Objetivos

1. Aplicar técnicas de transformación de datos: Realizar la limpieza, normalización y enriquecimiento de los datos en la base de datos de staging. La finalidad de estas transformaciones es mejorar la calidad de los datos y garantizar su adecuación a las necesidades del modelo principal, todo ello con el propósito de contar con información lista para ser utilizada en el Data Mart.
2. Diseñar y ejecutar consultas SQL o scripts de carga: Con el fin de llevar los datos transformados desde la base de datos de staging al Data Mart final.
3. Incorporar información adicional relevante: A medida que se van creando nuevas tablas en el modelo estrella.

Planteamiento del problema

El problema central radica en cómo transformar y cargar estos datos de manera eficiente y efectiva en un Data Mart, asegurando que estén listos para un análisis detallado y relevante. Los datos provienen de múltiples sistemas origen, cada uno con diferentes estructuras y formatos. Los datos deben ser transformados y enriquecidos para cumplir con los requisitos del modelo estrella. Para abordar estos desafíos, es esencial establecer un proceso robusto de ETL (Extract, Transform, Load) que incluya la extracción de datos de las bases de datos, asegurando que se cubren todos los aspectos necesarios para el análisis, la aplicación de técnicas de limpieza, normalización, enriquecimiento y la transferencia de los datos transformados desde la base de datos de staging al Data Mart final.

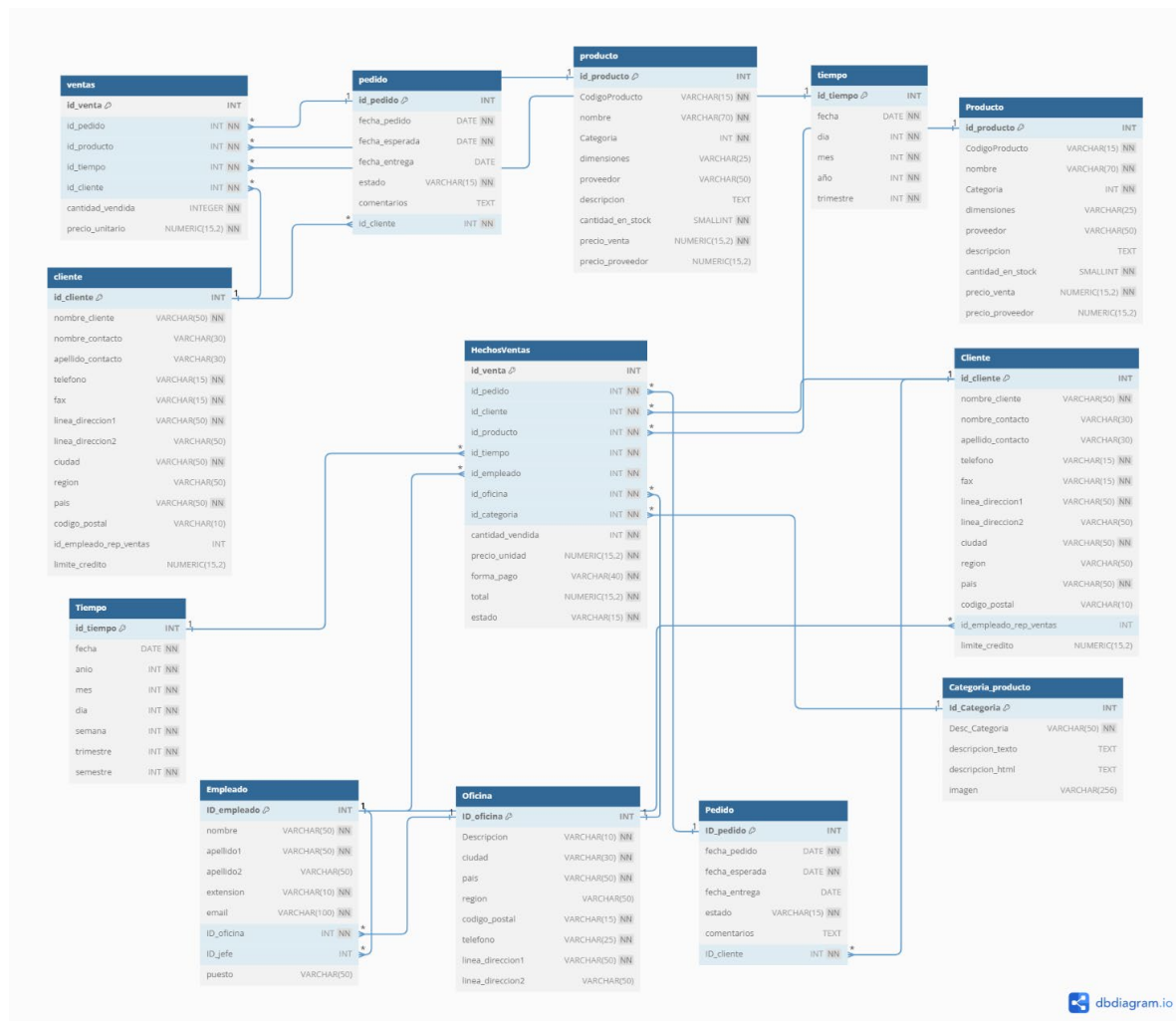
Análisis del problema

Para llevar a cabo las transformaciones, se emplean consultas SQL junto con herramientas de ETL. Mientras las consultas SQL posibilitan llevar a cabo operaciones de limpieza y transformación directamente en la base de datos de staging, las herramientas ETL son responsables tanto del manejo como la automatización eficiente del proceso de transformación de los datos. Durante esta etapa, es importante enfrentar diversos retos como asegurar la calidad de los datos, mantener consistencia en distintas fuentes y transformaciones, así como optimizar las consultas y procesos ETL para manejar grandes cantidades de información sin que esto impacte el rendimiento. Después, es necesario transferir los datos transformados desde la base de datos de staging hacia el Data Mart final. Una forma de lograr esto es a través del diseño de consultas SQL o scripts de carga que estén optimizados para insertar los datos en el Data Mart.

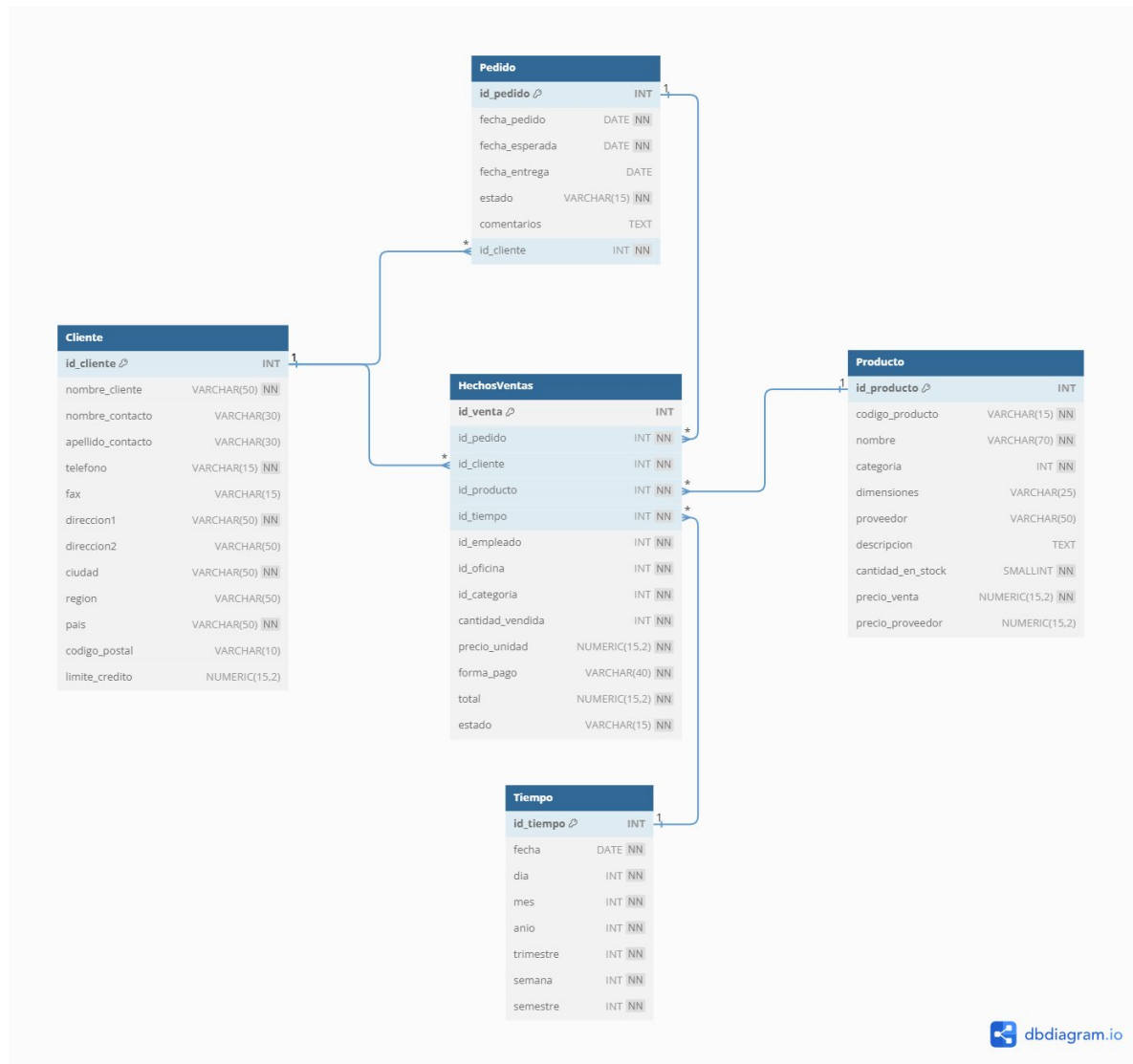
Propuesta de solución

Modelo de estrella:

MODELO ESTRELLA – PRIMERA ENTREGA

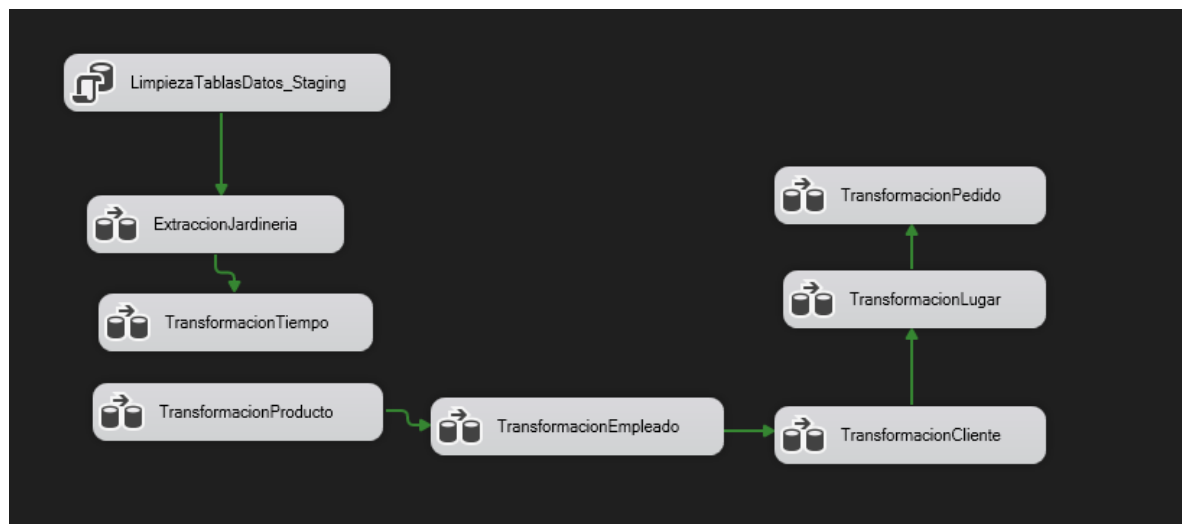


MODELO ESTRELLA – JARDINERIA CORREGIDA



Transformación de Datos según las Necesidades Analíticas

A continuación, se describe el proceso de transformación de datos para cada una de las dimensiones del modelo estrella:



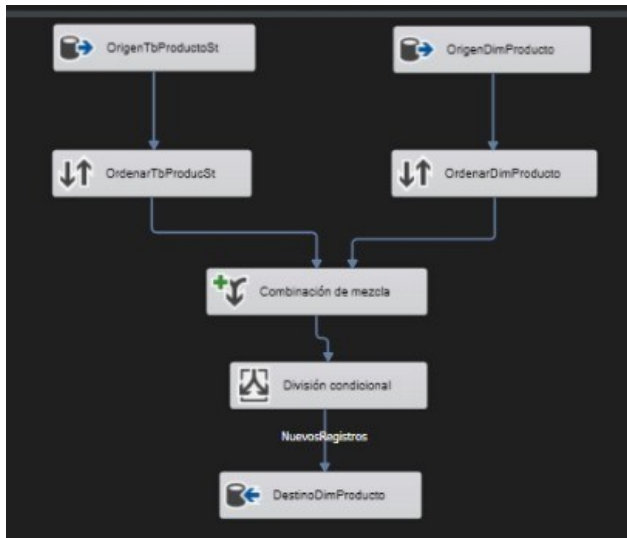
TransformacionProducto

Llamamos cada uno de los datos que harán parte de la dimensión producto de la data Mart.

```

select id_producto, id_producto_0, codigo_producto, nombre, COALESCE(NULLIF(LTRIM(RTRIM(dimensiones)), ''), 'N/A') as dimensiones,
nombre_proveedor, COALESCE(NULLIF(LTRIM(RTRIM(descripcion)), ''), 'N/A') as descripcion,
cantidad_en_stock, precio_unidad_proveedor, TbC.Desc_Categoria as nombre_categoria, TbC.descripcion_texto as descripcion_categoria
from TbProducto inner join TbCategoria TbC on Categoria = TbC.Id_Categoria_0
  
```

Esta declaración SQL se agrega al flujo de trabajo de la dimensión en ADO.NET usando OrigenTbProductoST. Antes de aplicar la combinación y mezcla con el OrigenDimProducto, se procede a ordenar estos registros por ID de producto. Luego, se realiza una división condicional para verificar si hay nuevos registros insertados utilizando la función isNull() en los IDs de productos obtenidos. DestinoDimProducto contará con la sentencia SQL para crear la tabla como dimensión (DimProducto), incluyendo todos los registros ordenados, normalizados y actualizados, además de las asignaciones predeterminadas. Se muestra de esta manera:



TransformacionEmpleado

Llamamos cada uno de los datos que harán parte de la dimensión producto de la data Mart.

```

select id_empleado, id_empleado_0, nombre,
apellido1, COALESCE(NULLIF(LTRIM(RTRIM(apellido2)), ''), 'N/A') as apellido2, extension, email,
Tb0.Descripcion as oficina, puesto from TbEmpleado TbE inner join TbOficina Tb0 on TbE.ID_oficina = Tb0.ID_oficina_0
  
```

Esta consulta en SQL forma parte de un proceso dentro del flujo de trabajo de la dimensión en ADO. Se genera la tabla DimEmpleado en DestinoDimEmpleado al combinar los registros de OrigenTbEmpleadoST, ordenándolos por ID de empleado y actualizándolos según corresponda con las asignaciones predeterminadas. Este proceso garantiza la consistencia y actualización de los datos en la base de datos junto con la información obtenida desde OrigenDimEmpleado.

Se muestra de esta manera:



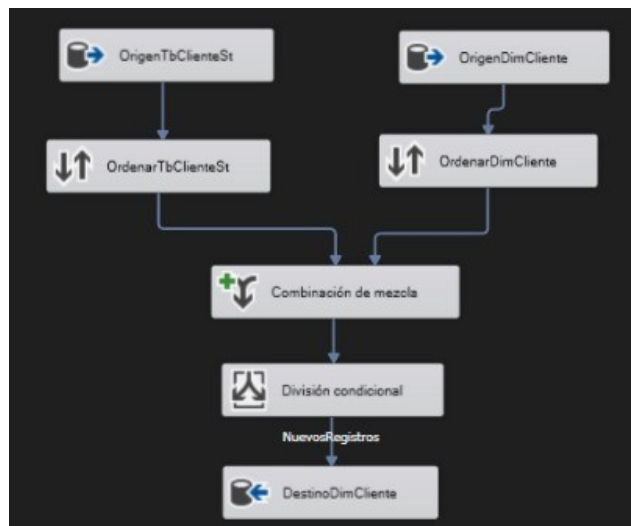
TransformacionCliente

Llamamos cada uno de los datos que harán parte de la dimensión producto de la data Mart.

```
select id_cliente, id_cliente_O,  
nombre_cliente, nombre_contacto, apellido_contacto,  
telefono, fax, linea_direccion1, COALESCE(NULLIF(LTRIM(RTRIM(linea_direccion2)), ''), 'N/A') as linea_direccion2,  
ciudad, COALESCE(NULLIF(LTRIM(RTRIM(region)), ''), 'N/A') as region from TbCliente
```

Se incorpora esta sentencia SQL en el flujo de trabajo de la dimensión mediante ADO. En NET OrigenTbClienteST, se eligen los registros de la tabla TbCliente. La sentencia recopila información detallada de los clientes y se fusiona con OrigenDimCliente. A continuación, se insertan en DestinoDimCliente para mantener actualizada la información ordenada de los registros. Esto garantiza que la base de datos tenga una completa consistencia en los detalles de cada cliente.

Se muestra de esta manera:



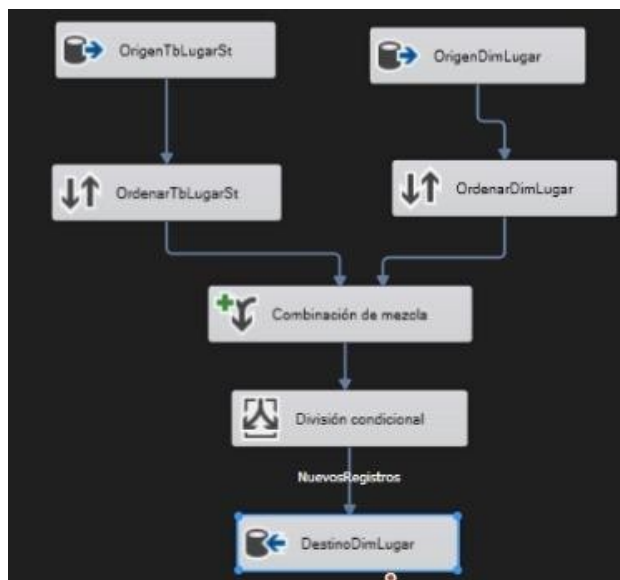
TransformaciónLugar

A continuación realizamos un llamado de cada uno de los datos que hacen parte de la dimensión Lugar en el DataMart, en este caso son: los campos de id_lugar, pais, region, ciudad, linea_direccion1.

```
select id_lugar, pais, region, ciudad, linea_direccion1 as direccion from TbLugar
```

La sentencia presentada se inserta dentro del flujo de trabajo de la dimensión **OrigenTbLugarSt**, en la cual los registros se ordenan por ID de lugar y se realiza la combinación con **OrigenDimLugar**, la consulta AS se está utilizando para renombrar la columna linea_direccion1 como dirección en el resultado de la consulta, esto quiere decir que los datos que se extraen dentro de la columna linea_direccion1 se mostrarán bajo la columna dirección.

Se muestra de esta manera:



TransformaciónTiempo

Se muestra de esta manera:



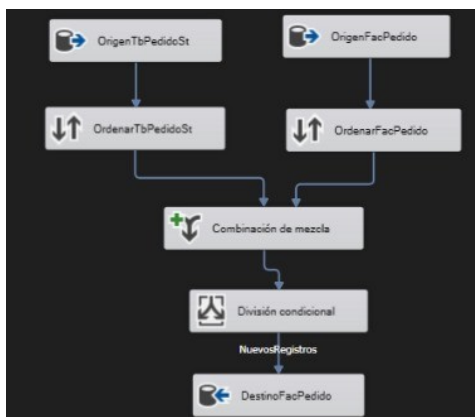
TransformacionPedido

Llamamos cada uno de los datos que harán parte de la dimensión producto escogido como Fact de la data Mart.

```
select TbP.id_pedido, id_pedido_0, DC.id_cliente, DP.id_producto, DT.id_tiempo, DE.id_empleado, DL.id_lugar,
TbD.cantidad, TbD.precio_unidad, TP.forma_pago, TP.total, estado
from TbPedido TbP inner join TbDetallePedido TbD on TbP.id_pedido_0 = TbD.ID_pedido
inner join DimProducto DP on TbD.ID_producto = DP.id_producto_0 inner join DimCliente DC on TbP.ID_cliente = DC.id_cliente_0
inner join TbCliente TbC on TbC.id_cliente_0 = TbP.ID_cliente
inner join DimTiempo DT on DT.fecha_pedido = TbP.fecha_pedido
inner join DimEmpleado DE on TbC.ID_empleado_rep_ventas = DE.id_empleado_0 inner join TbEmpleado TbE on TbE.id_empleado_0 = DE.id_empleado_0
inner join TbOficina TbO on TbO.ID_oficina_0 = TbE.ID_oficina inner join DimLugar DL on DL.direccion = TbO.linea_direccion1
inner join TbPago TP on TP.ID_cliente = DC.id_cliente_0
```

Al combinar datos de múltiples tablas relacionadas, esta consulta SQL recupera información detallada sobre los pedidos. La opción INNER JOIN se emplea para combinar la información basándose en campos de identificación compartidos, por ejemplo, el número de pedido, cliente, producto, fecha y hora del pedido, empleado a cargo y ubicación. El sistema brinda una visión detallada de cada pedido, mostrando información sobre la cantidad solicitada, el precio unitario del producto elegido por el cliente y la forma de pago utilizada.

Se muestra de esta manera:



Carga de registros en el Data Mart final:

Evidencia de los datos transformados en la dimensión Producto:

```
/****** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP (1000) [id_producto]
, [id_producto_0]
, [codigo_producto]
, [nombre]
, [dimensiones]
, [nombre_proveedor]
, [descripcion]
, [cantidad_en_stock]
, [precio_unidad_proveedor]
, [nombre_categoria]
, [descripcion_categoria]
FROM [Staging-Jardineria].[dbo].[DimProducto]
```

	codigo_producto	nombre	dimensiones	nombre_proveedor	descripcion	cantidad_en_stock	precio_unidad_proveedor	nombre_categoria	descripcion_categoria
172	OR-138	Camelia Blanco, Chrysler Rojo...	N/A	Viveros EL OASIS	N/A	100	3.00	Omamentales	Plantas vistosas para la decoración del
173	OR-139	Landora Amarillo, Rose Gaujar...	N/A	Viveros EL OASIS	N/A	100	3.00	Omamentales	Plantas vistosas para la decoración del
174	OR-140	Kordes Perfect bicolor rojo-am...	N/A	Viveros EL OASIS	N/A	100	3.00	Omamentales	Plantas vistosas para la decoración del
175	OR-141	Pitmini rojo	N/A	Viveros EL OASIS	N/A	100	3.00	Omamentales	Plantas vistosas para la decoración del
176	OR-142	Solanum Jazminoides	150-160	Viveros EL OASIS	N/A	100	1.00	Omamentales	Plantas vistosas para la decoración del
177	OR-143	Wisteria Sinensis azul, rosa, bl...	N/A	Viveros EL OASIS	N/A	100	7.00	Omamentales	Plantas vistosas para la decoración del
178	OR-144	Wisteria Sinensis INJERTADAS...	140-150	Viveros EL OASIS	N/A	100	9.00	Omamentales	Plantas vistosas para la decoración del
179	OR-145	Bougainvillea Sandieriana Tutor	80-100	Viveros EL OASIS	N/A	100	1.00	Omamentales	Plantas vistosas para la decoración del
180	OR-146	Bougainvillea Sandieriana Tutor	125-150	Viveros EL OASIS	N/A	100	3.00	Omamentales	Plantas vistosas para la decoración del
181	OR-157	Acer Pseudoplatanus	200-225	Viveros EL OASIS	N/A	100	8.00	Omamentales	Plantas vistosas para la decoración del
182	OR-158	Brachychiton Acerifolius	200-225	Viveros EL OASIS	N/A	100	4.00	Omamentales	Plantas vistosas para la decoración del
183	OR-159	Brachychiton Discolor	200-225	Viveros EL OASIS	N/A	100	4.00	Omamentales	Plantas vistosas para la decoración del
184	OR-160	Brachychiton Rupestris	170-200	Viveros EL OASIS	N/A	100	8.00	Omamentales	Plantas vistosas para la decoración del
185	OR-161	Cassia Corimbosa	200-225	Viveros EL OASIS	N/A	100	4.00	Omamentales	Plantas vistosas para la decoración del
186	OR-162	Cassia Corimbosa	200-225	Viveros EL OASIS	N/A	100	8.00	Omamentales	Plantas vistosas para la decoración del
187	OR-163	Chitaloa Summer Bells	200-225	Viveros EL OASIS	N/A	80	8.00	Omamentales	Plantas vistosas para la decoración del

Evidencia de los datos transformados en la dimensión Empleado:

```
/****** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP (1000) [id_empleado]
, [id_empleado_0]
, [nombre]
, [apellido1]
, [apellido2]
, [extension]
, [email]
, [oficina]
, [puesto]
FROM [Staging-Jardineria].[dbo].[DimEmpleado]
```

	id_empleado	id_empleado_0	nombre	apellido1	apellido2	extension	email	oficina	puesto
1	1	1	Marcos	Magaña	Perez	3897	marcos@jardineria.es	TAL-ES	Director General
2	2	2	Ruben	López	Martinez	2899	rioiez@jardineria.es	TAL-ES	Subdirector Marketing
3	3	3	Alberto	Soria	Carrasco	2837	asoria@jardineria.es	TAL-ES	Subdirector Ventas
4	4	4	Maria	Solis	Jerez	2847	msolis@jardineria.es	TAL-ES	Secretaria
5	5	5	Felipe	Rosas	Marquez	2844	frosas@jardineria.es	TAL-ES	Representante Ventas
6	6	6	Juan Carlos	Ortiz	Serrano	2845	cortiz@jardineria.es	TAL-ES	Representante Ventas
7	7	7	Carlos	Soria	Jimenez	2444	csoria@jardineria.es	MAD-ES	Director Oficina
8	8	8	Mariano	López	Murcia	2442	mlopez@jardineria.es	MAD-ES	Representante Ventas
9	9	9	Lucio	Campoamor	Martin	2442	lcampoamor@jardineria.es	MAD-ES	Representante Ventas
10	10	10	Hilario	Rodriguez	Huertas	2444	hrodriguez@jardineria.es	MAD-ES	Representante Ventas
11	11	11	Emmanuel	Magaña	Perez	2518	manu@jardineria.es	BCN-ES	Director Oficina
12	12	12	José Manuel	Martinez	De la Osa	2519	jmmart@hotmail.es	BCN-ES	Representante Ventas
13	13	13	David	Palma	Aceituno	2519	dpalma@jardineria.es	BCN-ES	Representante Ventas
14	14	14	Oscar	Palma	Aceituno	2519	opalma@jardineria.es	BCN-ES	Representante Ventas
15	15	15	Francois	Fignon	N/A	9981	ffignon@gardening.com	PAR-FR	Director Oficina
16	16	16	Lionel	Narvaez	N/A	9982	lnarvaez@gardening.com	PAR-FR	Representante Ventas
17	17	17	Laurent	Serra	N/A	9982	lserra@gardening.com	PAR-FR	Representante Ventas

Evidencia de los datos transformados en la dimensión Cliente:

***** Script para el comando SelectTopNRows de SSMS *****

```

SELECT TOP (1000) [id_cliente]
, [id_cliente_O]
, [nombre_cliente]
, [nombre_contacto]
, [apellido_contacto]
, [telefono]
, [fax]
, [linea_direccion1]
, [linea_direccion2]
, [ciudad]
, [region]
FROM [Staging-Jardineria].[dbo].[DimCliente]

```

100 %

Resultados Mensajes

	id_cliente	id_cliente_O	nombre_cliente	nombre_contacto	apellido_contacto	telefono	fax	linea_direccion1	linea_direccion2	ciudad	region
1	1	1	GoldFish Garden	Daniel G	GoldFish	5556901745	5556901746	False Street 52 2 A	NA	San Francisco	NA
2	2	2	Gardening Associates	Anne	Wright	5557410345	5557410346	Wall-e Avenue	NA	Miami	Miami
3	3	3	Gendo Valley	Link	Flaute	5552323129	5552323128	Oaks Avenue n°22	NA	New York	NA
4	4	4	Tendo Garden	Akane	Tendo	55591233210	55591233211	Null Street n°69	NA	Miami	NA
5	5	5	Lasas S.A.	Antonio	Lasas	34916540145	34914851312	C/Leganes 15	NA	Fuenlabrada	Madrid
6	6	6	Beragua	Jose	Bermejo	654867321	916549672	C/pintor segundo	Getafe	Madrid	Madrid
7	7	7	Club Golf Puerta del hierro	Peco	Lopez	624566810	919535678	C/inesio delgado	Madrid	Madrid	Madrid
8	8	8	Naturagua	Guillermo	Rengifo	689234750	916428956	C/majadahonda	Boadilla	Madrid	Madrid
9	9	9	DaraDistribuciones	David	Serrano	675598001	916421756	C/azores	Fuenlabrada	Madrid	Madrid
10	10	10	Madriña de riegos	Jose	Tacaño	655883045	916689215	C/Lagañas	Fuenlabrada	Madrid	Madrid
11	11	11	Lasas S.A.	Antonio	Lasas	34916540145	34914851312	C/Leganes 15	NA	Fuenlabrada	Madrid
12	12	12	Camunas Jardines S.L.	Pedro	Camunas	34914873241	34914871541	C/Virgenes 45	C/Princesas 2 11B	San Lorenzo del Escorial	Madrid
13	13	13	Dardena S.A.	Juan	Rodriguez	34912453217	34912484764	C/Nueva York 74	NA	Madrid	Madrid
14	14	14	Jardin de Flores	Javier	Villar	654865643	914538776	C/ Oña 34	NA	Madrid	Madrid
15	15	15	Flores Marivi	Maria	Rodriguez	666555444	912458657	C/Leganes24	NA	Fuenlabrada	Madrid
16	16	16	Flowers S.A	Beatriz	Fernandez	698754159	978453216	C/Luis Salquillo4	NA	Montomes del valles	Barcelona
17	17	17	Naturajardin	Victoria	Cruz	612343529	916548735	Plaza Maquilón 15	NA	Madrid	Madrid

Evidencia de los datos transformados en la dimensión Lugar:

***** Script para el comando SelectTopNRows de SSMS *****

```

SELECT TOP (1000) [id_lugar]
, [pais]
, [region]
, [ciudad]
, [direccion]
FROM [Staging-Jardineria].[dbo].[DimLugar]

```

%

Resultados Mensajes

	id_lugar	pais	region	ciudad	direccion
1	1	EEUU	CA	San Francisco	100 Market Street
2	2	EEUU	MA	Boston	1550 Court Place
3	3	Francia	EMEA	Paris	29 Rue Jouffroy d'abbans
4	4	Japón	Chiyoda-Ku	Tokyo	4-1 Kioicho
5	5	Australia	APAC	Sydney	5-11 Wentworth Avenue
6	6	Inglaterra	EMEA	Londres	52 Old Broad Street
7	7	España	Barcelona	Barcelona	Avenida Diagonal, 38
8	8	España	Madrid	Madrid	Bulevar Indalecio Prieto, 32
9	9	España	Castilla-LaMancha	Talavera de la Reina	Francisco Aguirre, 32

Evidencia de los datos transformados en la dimensión Tiempo:

Derived Column Name	Derived Column	Expression	Data Type	U
fecha	<add as new column>	(DT_WSTR,2)DAY(fecha_pedido) + "-" + (DT_...	cadena Unicode [DT_...	1
año	<add as new column>	YEAR(fecha_pedido)	entero de cuatro byt...	
mes	<add as new column>	MONTH(fecha_pedido)	entero de cuatro byt...	
día	<add as new column>	DAY(fecha_pedido)	entero de cuatro byt...	
semana	<add as new column>	DATEPART("dw",fecha_pedido)	entero de cuatro byt...	
semestre	<add as new column>	(DATEPART("Month",fecha_pedido) + 5) / 6	entero de cuatro byt...	
bimestre	<add as new column>	(DATEPART("Month",fecha_pedido) + 1) / 2	entero de cuatro byt...	

Evidencia de los datos transformados en la dimensión Pedido:

```

/***** Script para el comando SelectTopNRows de SSMS *****/
SELECT TOP (1000) [id_pedido]
, [id_pedido_0]
, [id_cliente]
, [id_producto]
, [id_tiempo]
, [id_empleado]
, [id_lugar]
, [cantidad]
, [precio_unidad]
, [forma_pago]
, [total]
, [estado]
FROM [Staging-Jardineria].[dbo].[FacPedido]

```

100 %

Resultados Mensajes

	id_pedido	id_pedido_0	id_cliente	id_producto	id_tiempo	id_empleado	id_lugar	cantidad	precio_unidad	forma_pago	total	estado
1	1	1	5	87	1	8	8	10	70.00	Transferencia	23794.00	Entregado
2	1	1	5	151	1	8	8	40	4.00	Transferencia	23794.00	Entregado
3	1	1	5	165	1	8	8	25	4.00	Transferencia	23794.00	Entregado
4	1	1	5	265	1	8	8	15	19.00	Transferencia	23794.00	Entregado
5	1	1	5	276	1	8	8	23	14.00	Transferencia	23794.00	Entregado
6	2	2	5	271	8	8	8	5	462.00	Transferencia	23794.00	Entregado
7	2	2	5	251	8	8	8	67	64.00	Transferencia	23794.00	Entregado
8	2	2	5	183	8	8	8	12	6.00	Transferencia	23794.00	Entregado
9	2	2	5	58	8	8	8	7	8.00	Transferencia	23794.00	Entregado
10	2	2	5	164	8	8	8	50	4.00	Transferencia	23794.00	Entregado
11	2	2	5	57	8	8	8	3	29.00	Transferencia	23794.00	Entregado
12	2	2	5	165	8	8	8	20	5.00	Transferencia	23794.00	Entregado
13	3	3	5	241	14	8	8	15	65.00	Transferencia	23794.00	Rechazado
14	3	3	5	237	14	8	8	30	266.00	Transferencia	23794.00	Rechazado
15	3	3	5	242	14	8	8	24	25.00	Transferencia	23794.00	Rechazado
16	3	3	5	146	14	8	8	32	5.00	Transferencia	23794.00	Rechazado
17	3	3	5	66	14	8	8	120	9.00	Transferencia	23794.00	Rechazado

Conclusiones

- **Construcción y Integración de la Base de Datos Staging:** Es fundamental establecer y luego fusionar una base de datos staging con un Data Mart para poder administrar los datos de manera eficiente. En este proceso se realiza la revisión del modelo estrella, se extraen datos de diferentes fuentes, se aplican transformaciones para mejorar la calidad y consistencia de los datos, y por último se cargan estos datos transformados en el Data Mart. Los retos principales a enfrentar son garantizar la calidad y uniformidad de los datos, así como mejorar el procedimiento ETL para trabajar con grandes cantidades de información sin afectar el rendimiento.
- **Mejora en la Gestión de Datos:** Lograr con éxito la implementación de una base de datos staging y su integración en un Data Mart brinda una sólida plataforma para realizar análisis detallados y mejorar la gestión de datos. Este logra cumplir los objetivos planteados y establece un alto estándar para proyectos de datos futuros, garantizando que la información esté disponible y sea confiable al momento de tomar decisiones estratégicas.