



IU Digital
de Antioquia



Proceso de transformación de datos y carga en el data mart final

Ardys Diaz Hurtado

Institución Universitaria Digital de Antioquia
Ingeniería de Software y Datos

Quibdó – Chocó
2024

Introducción

En el competitivo mundo empresarial actual, la capacidad de gestionar y analizar datos de manera eficiente es crucial para la toma de decisiones estratégicas. Este documento presenta el desarrollo de un proceso de transformación y carga de datos, desde la base de datos origen, pasando por una base de datos intermedia (Staging), hasta el data mart final. El objetivo es garantizar que los datos estén listos para el análisis, incluyendo la identificación de los productos más vendidos y otros análisis importantes para el negocio.

Objetivos

El objetivo principal de este proyecto es implementar un proceso ETL (Extracción, Transformación y Carga) que prepare adecuadamente los datos para su análisis. Los objetivos específicos son:

1. **Extracción:** Obtener datos relevantes de la base de datos origen.
2. **Transformación:** Limpiar, transformar y estructurar los datos en la base de datos Staging.
3. **Carga:** Transferir los datos transformados al data mart final, asegurando su preparación para el análisis.
4. **Análisis:** Facilitar la identificación de los productos más vendidos y realizar otros análisis que apoyen la toma de decisiones empresariales.

Planteamiento del Problema

La empresa necesita un proceso robusto para gestionar y preparar sus datos para análisis. Actualmente, enfrenta desafíos como:

1. **Integridad de Datos:** Falta de un proceso centralizado que asegure la coherencia y calidad de los datos.
2. **Eficiencia:** Procesos manuales propensos a errores y que consumen mucho tiempo.
3. **Capacidad de Análisis:** Dificultades para realizar análisis detallados debido a datos no estructurados o mal organizados.

Análisis del Problema

El análisis de la situación actual revela la necesidad de un proceso ETL bien estructurado que permita:

1. **Centralización de Datos:** Consolidar datos dispersos en una única base de datos Staging.
2. **Transformación de Datos:** Aplicar reglas de limpieza y transformación para asegurar que los datos sean consistentes y útiles.
3. **Preparación para el Análisis:** Transferir los datos transformados a un data mart final donde se puedan realizar análisis detallados.

Descripción del Proceso de ETL

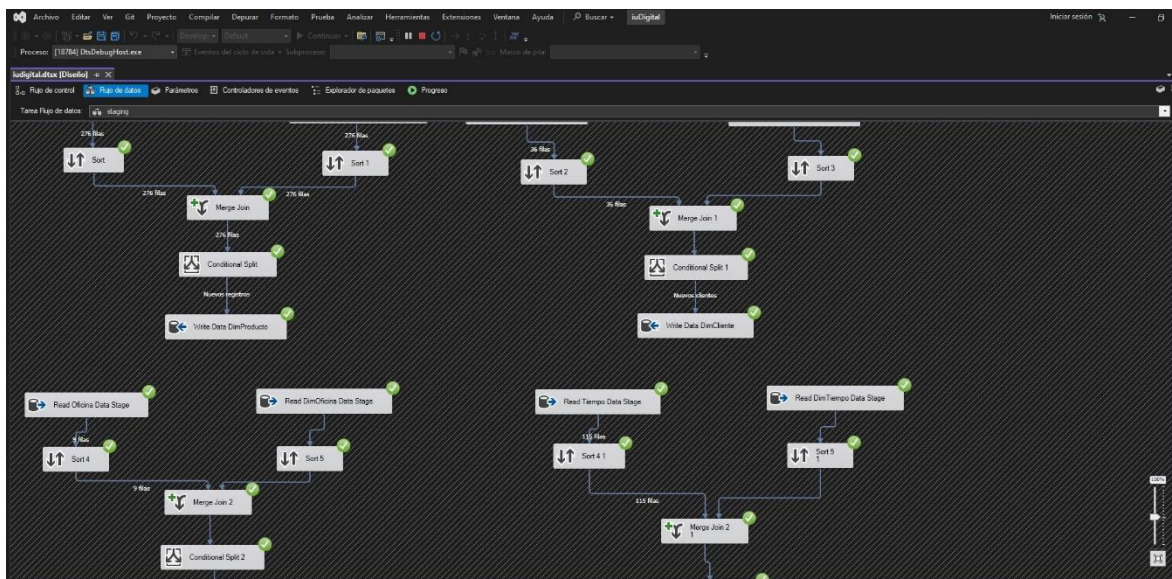
1. **Extracción de Datos:** Se utilizaron consultas SQL para extraer datos relevantes de la base de datos origen, seleccionando información crítica necesaria para el análisis posterior.
2. **Transformación de Datos:** En la base de datos Staging, se realizaron actividades como:
 - **Limpieza de Datos:** Eliminación de duplicados y corrección de inconsistencias.
 - **Transformación:** Conversión de datos a formatos adecuados, creación de nuevas variables derivadas y aplicación de reglas de negocio.
3. **Carga de Datos:** Los datos transformados se cargaron en el data mart final. Este paso aseguró que la información estuviera organizada de manera que facilite su análisis.

Conclusiones

El desarrollo de un proceso de transformación y carga de datos desde la base de datos origen hasta el data mart final ha mejorado significativamente la capacidad de la empresa para analizar sus datos y tomar decisiones informadas. La implementación de una base de datos Staging como paso intermedio ha permitido

asegurar la calidad y consistencia de los datos, facilitando la identificación de los productos más vendidos y otros análisis críticos. En resumen, este proyecto ha optimizado la gestión de datos y ha proporcionado una infraestructura sólida para el análisis empresarial.

Anexos



SQL Query Log - 177.93.58.58 (sa (100)) - Microsoft SQL Server Management Studio

Quick Launch (Ctrl-Q)

Object Explorer

SQL Query Log - 177...mensaDIM (sa (100))

SQL Query Log - 177...mensaDIM (sa (100))

100 % Results

ID	producto	nombre	categoria	proveedor
1	Sierra de Poda 400MM	Herramientas	HiperGarden Tools	
2	Pala	Herramientas	HiperGarden Tools	
3	Plástico de Jardín	Herramientas	HiperGarden Tools	
4	Escalón	Herramientas	HiperGarden Tools	
5	Agallas	Armonías	Murcia Seasons	
6	Levándula Dentada	Armonías	Murcia Seasons	
7	Majonea	Armonías	Murcia Seasons	
8	Melissa	Armonías	Murcia Seasons	
9	Morha Salvia	Armonías	Murcia Seasons	
10	Petalatum Vinterae (Peregr)	Armonías	Murcia Seasons	
11	Salsa Mla	Armonías	Murcia Seasons	
12	Thymus Citriodora (Tambú Indio)	Armonías	Murcia Seasons	
13	Thymus Vulgaris	Armonías	Murcia Seasons	
14	Santolina Chamaecypariss	Armonías	Murcia Seasons	
15	Euphorbia Chrysops Mla	Frutales	Frutales Talavera S.A	
16	Limonero 2 años injerto	Frutales	NaranjaValencianas.com	
17	Nectama	Frutales	Frutales Talavera S.A	
18	Nogal	Frutales	Frutales Talavera S.A	
19	Oliva Oliva	Frutales	Frutales Talavera S.A	
20	Oliva Oliva	Frutales	Frutales Talavera S.A	
21	Oliva Oliva	Frutales	Frutales Talavera S.A	
22	Oliva Oliva	Frutales	Frutales Talavera S.A	
23	Peral	Frutales	Frutales Talavera S.A	
24	Peral	Frutales	Frutales Talavera S.A	
25	Peral	Frutales	Frutales Talavera S.A	
26	Limonero 30-40	Frutales	NaranjaValencianas.com	
27	Kumquat	Frutales	NaranjaValencianas.com	
28	Kumquat EXTRA con FRUITA	Frutales	NaranjaValencianas.com	
29	Calamondin Mla	Frutales	Frutales Talavera S.A	
30	Calamondin Copa	Frutales	Frutales Talavera S.A	
31	Calamondin Copa EXTRA Co...	Frutales	Frutales Talavera S.A	
32	Rosal bajo 1M En macetas...	Frutales	Frutales Talavera S.A	
33	RISAL TRESPASADOR	Frutales	Frutales Talavera S.A	
34	Camelia Blanco, Chrysler Rd...	Frutales	NaranjaValencianas.com	
35	Naranga- Plantón joven 1 año...	Frutales	NaranjaValencianas.com	
36	Laridona Ananillo, Rosa Gra...	Frutales	Frutales Talavera S.A	
37	Rosita Perfect local roje...	Frutales	Frutales Talavera S.A	
38	Platan rojo	Frutales	Frutales Talavera S.A	
39	Rosal copa	Frutales	Frutales Talavera S.A	
40	Robinsonian Cultivo	Frutales	Melocotones de Cien...	

Query executed successfully

177.93.58.58 (12.0 RTM) sa (100) jardenariaDIM 00:00:00 276 rows

Bibliografía

- Jordán Sierra, J. S. (2022). Arquitectura ETL, técnicas de procesamiento y análisis de datos para hallar relaciones entre la producción minera y los índices de educación y pobreza de Colombia.
- Manso Millán, J. A. (2021). Retos de la Migración de Datos hacia un Data Warehouse.