

EA 2. CREACIÓN DE UNA BASE DE DATOS DE STAGING

Carlos Alberto Ibarguen Urrutia CC1017132063

Carlos Alberto Asprilla CC: 1017166777

Yeferson Ibargüen Urrutia CC : 1030450304

Anderson Andres Castrillos CC 70330170

Grupo # 11

Ingeniería de Software y Datos

Institución Universitaria Digital de Antioquia

Semestre 7º

Bases de Datos II

Víctor Hugo Mercado

14/03/2025

ÍNDICE

EA 2. CREACIÓN DE UNA BASE DE DATOS DE STAGING	1
INTRODUCCIÓN	2
OBJETIVOS	4
PLANTEAMIENTO DEL PROBLEMA Y ANÁLISIS	4
PROPUESTA DE LA SOLUCIÓN ;Error! Marcador no	
definido. CORRECCIONES A LA ENTREGA 1 ;Error! Marcador	
no definido.	
DESCRIPCIÓN DEL ANÁLISIS REALIZADO A LOS DATOS	5
ANEXOS	7
CONCLUSIONES	7
REFERENCIAS	9

CONCLUSIÓN

El proceso de crear un modelo estrella y una base de datos de staging es esencial para manejar y procesar grandes volúmenes de datos de manera efectiva. Al integrar el modelo estrella, conseguimos organizar los datos de manera estructurada y clara, lo que facilita tanto su análisis como la toma de decisiones. Este enfoque es particularmente útil cuando se trabaja con bases de datos complejas, como la de Jardinería, donde la información puede estar dispersa y no siempre es de fácil acceso.

En primer lugar, el análisis de los datos de la base de datos Jardinería es un paso fundamental antes de la creación de la base de datos de staging. Este proceso nos permite identificar qué información es relevante para el propósito de nuestra nueva estructura de datos. Solo aquellos datos esenciales serán trasladados a la base de datos de staging, garantizando así que no se introduzcan datos innecesarios que podrían complicar el posterior procesamiento y análisis.

Luego, una vez seleccionados los datos, diseñamos las tablas de la base de datos de staging. Este diseño no solo debe ser eficiente en términos de almacenamiento, sino también en términos de accesibilidad y facilidad de transformación. Un diseño adecuado de la base de datos de staging asegura que las consultas SQL que realicemos para transferir los datos desde la base de datos Jardinería sean eficaces y puedan ejecutarse sin problemas. Es crucial validar estas consultas para asegurar que los datos se transfieren correctamente, evitando errores en el almacenamiento y asegurando que la base de datos de staging esté lista para su uso posterior en el análisis de datos.

Finalmente, todo el proceso, desde la selección de los datos relevantes hasta la validación de las consultas SQL, es un paso crucial hacia la creación de un entorno de datos eficiente y bien organizado. La base de datos de staging actúa como un punto intermedio donde los datos se almacenan temporalmente antes de ser procesados y utilizados en análisis más complejos, lo que optimiza la toma de decisiones y la generación de reportes en el modelo estrella.

En resumen, la creación de un modelo estrella y una base de datos de staging bien diseñada es una estrategia eficaz para manejar datos de manera estructurada y optimizada, lo que no solo mejora la calidad del análisis de datos, sino que también facilita la toma de decisiones informadas y la eficiencia operativa.

OBJETIVOS

El objetivo principal de diseñar un Modelo Estrella para la base de datos de Jardinería es facilitar la extracción de información clave que apoye la toma de decisiones estratégicas. Por ejemplo, determinar el año con mayores ventas, identificar el producto de mayor demanda y analizar la categoría con mayor cantidad de productos. Este modelo, al ser estructurado de manera eficiente, también busca mejorar la agilidad en la manipulación de la base de datos, permitiendo la integración sencilla de nuevas dimensiones y datos conforme evolucionan las necesidades comerciales y de análisis.

El propósito fundamental de esta actividad es desarrollar consultas SQL que permitan crear una base de datos staging a partir de la estructura actual de la base de datos Jardinería. La base de datos staging proporcionará una plataforma optimizada y organizada para facilitar el análisis de datos, permitiendo realizar consultas más rápidas y eficientes. Esto permitirá identificar patrones, tendencias y relaciones significativas, lo que será crucial para impulsar decisiones informadas en diversas áreas de la empresa, como las estrategias comerciales y de ventas.

PLANTEAMIENTO DEL PROBLEMA Y ANÁLISIS

La empresa Jardinería busca obtener información clave sobre sus operaciones, como el producto más vendido, la categoría con mayor cantidad de productos y el año con mayor volumen de ventas. Sin embargo, la actual estructura de su base de datos no facilita este análisis detallado.

El desafío principal radica en la necesidad de construir una base de datos staging basada en el modelo relacional de la base de datos Jardinería. Esta base de datos intermedia permitirá la preparación y transformación de datos para su análisis final.

Para lograr este objetivo, es esencial identificar los datos relevantes y diseñar una estructura de tablas adecuada que garantice la integridad de los datos. Además, se deben construir consultas SQL eficientes para extraer y cargar los datos en la base de datos staging.

Una vez transferidos los datos, se debe validar su correcto almacenamiento en la base de datos staging y establecer medidas de respaldo para garantizar la seguridad de la información.

DESCRIPCIÓN DEL ANÁLISIS REALIZADO A LOS DATOS

Para realizar el análisis de los datos de Jardinería y cómo los trasladamos a la base de datos Staging, primero tratamos de comprender la estructura de los datos y los requisitos que nos

pedía la actividad. Basándonos en las tablas y consultas para lograr el modelo estrella, hemos realizado el siguiente análisis:

Estructura de los datos de Jardinería: Los datos de Jardinería contienen información sobre oficinas, empleados, clientes, pedidos, productos, detalles de pedidos entre otros. Las tablas contienen información sobre los productos vendidos, las cantidades vendidas, los precios y los clientes que compraron los productos entre muchos más datos para analizar.

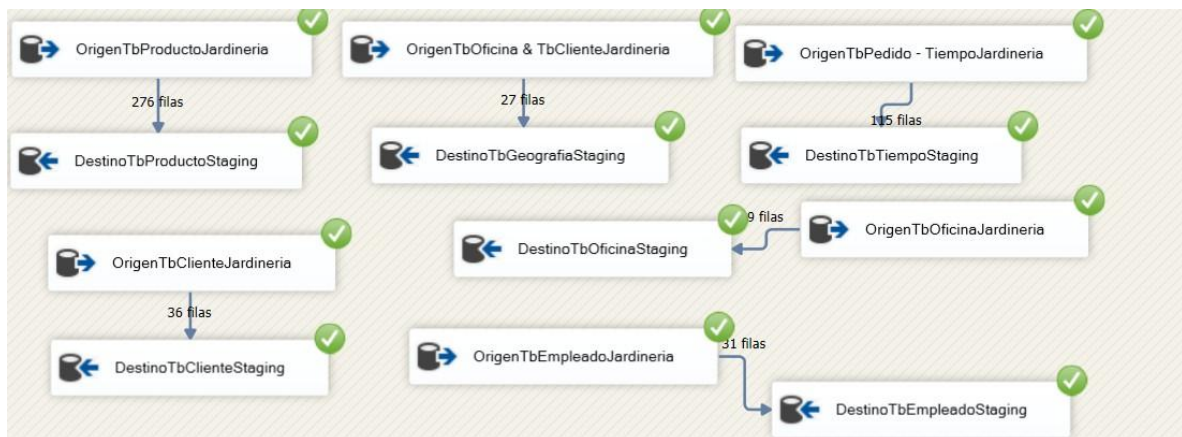
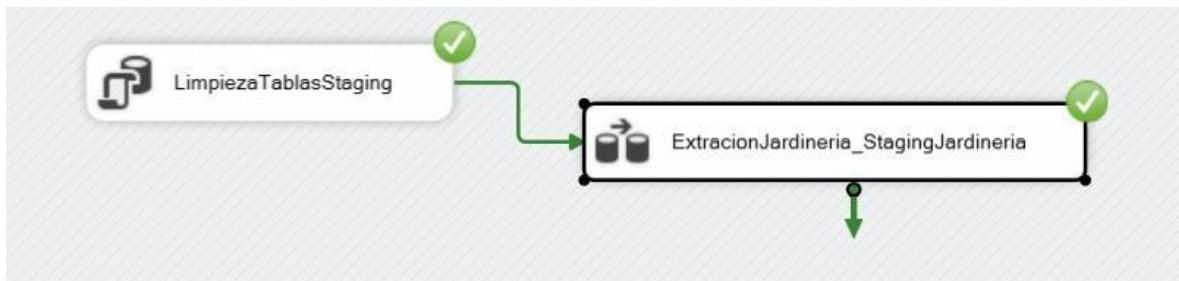
Análisis de los datos y relaciones: Las relaciones entre las tablas permiten realizar consultas que combinan datos de varias tablas para obtener información más completa sobre los clientes, pedidos, productos entre otros. Por ejemplo, se puede consultar la base de datos para encontrar nuestro principal objetivo “El año en el que hubo más ventas el producto más vendido y la categoría con mayor cantidad de productos”. Así como estos otros ejemplos:

- Los clientes que compraron más productos en un determinado período de tiempo.
- Los productos más vendidos en una determinada región.
- Las categorías de productos que generaron más ingresos.

Proceso de traslado a la base de datos Staging: El proceso de traslado de los datos de Jardinería a la base de datos Staging implicó los siguientes pasos:

1. **Extracción de datos:** Se utilizaron consultas SQL para extraer datos de las tablas de origen de la base de datos de Jardinería.
2. **Transformación de datos:** Los datos se limpiaron, enriquecieron y modificaron según los requisitos de la actividad. También se realizaron transformaciones para garantizar que los datos cumplieran con ciertas reglas de integridad, como la validación de formatos de fecha o la estandarización de nombres y direcciones.
3. **Carga de datos:** Los datos transformados se cargaron en la base de datos Staging.

Capturas de Pantalla como prueba:



ANEXOS

Los anexos estarán junto con este documento en el archivo comprimido

CONCLUSIONES

La construcción de la base de datos staging representó un desafío integral que implicó la extracción de datos relevantes de la base de datos original de Jardinería, su transformación conforme a los requisitos del proyecto y, finalmente, su carga en la nueva base de datos de staging. Este proceso no solo fue crucial para simplificar el análisis de los datos, sino que también permitió crear una estructura organizada y optimizada que facilitaría la toma de decisiones estratégicas en la empresa. Al tener los datos organizados de una manera eficiente, las consultas posteriores se realizarán de forma más rápida y precisa, lo que será esencial para el análisis y la interpretación efectiva de la información.

La creación de una base de datos staging, como paso intermedio entre la base de datos original y los modelos de análisis, es fundamental para la correcta preparación de los datos antes de su procesamiento en modelos más complejos, como el modelo estrella. Este modelo, basado en una estructura dimensional, simplifica la complejidad de los datos y proporciona una visión clara de las relaciones entre las diferentes dimensiones y medidas. Su implementación permite realizar análisis multidimensionales de manera eficiente, facilitando la identificación de patrones y tendencias que de otro modo serían difíciles de observar.

A lo largo del proceso, también fue evidente la importancia de una correcta estructuración de las tablas, ya que esto garantiza que los datos sean accesibles, coherentes y, lo más importante, relevantes para el análisis que se desea realizar. La relación entre las tablas debe ser cuidadosamente diseñada, ya que una estructura mal definida podría llevar a errores de interpretación o a la falta de información crucial. La base de datos staging debe ser vista como una etapa de preparación, donde los datos deben ser limpiados, transformados y organizados de forma adecuada para garantizar que las fases posteriores de análisis sean lo más eficientes y efectivas posible.

Además, el trabajo de transformación de datos en esta base de datos intermedia es una pieza clave para garantizar la calidad de la información que se utilizará en el análisis final. La precisión en la extracción y transformación de los datos es esencial para que los resultados de las consultas sean válidos y confiables. Un detalle como la correcta agregación de las ventas por producto o la correcta identificación de categorías de productos puede marcar una gran diferencia en los resultados finales y en la capacidad de tomar decisiones basadas en datos sólidos.

Este proceso de análisis exhaustivo de los datos y la comprensión profunda de las relaciones entre las tablas es lo que garantiza que la base de datos staging contenga información no solo estructurada, sino también relevante y útil. La correcta organización de la base de datos no solo facilita el acceso a la información, sino que también permite a los

usuarios ejecutar consultas complejas de forma más eficiente, sin perder tiempo en procesos de búsqueda y recuperación de datos.

REFERENCIAS

PowerData, R. (2024). *Staging: la salvaguarda de los procesos ETL*. Retrieved from <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/312643/Staginglasalvaguarda-de-los-procesos-ETL>

Alfaro, F. R. (2021). *El modelo de estrella. El pilar fundamental del Business Intelligence*. Retrieved from <https://datamanagement.es/2019/06/27/business-intelligence-modeloestrella/>

Cardona, G. (2023). *Qué es el Modelo Estrella: Aglaia Consulting*. Retrieved from <https://aglaia.es/blog/power-bi/que-es-el-modelo-estrella-en-power-bi/>

Zerpa, H., Garcia, R., & Izquierdo, H. (2020). Datamart basado en el modelo estrella para la implementación de indicadores clave de desempeño como salida del Big Data. *Universidad Ciencia y Tecnología*, 24(102), 47-54.

¿Qué es un data mart? (2023). Retrieved from <https://www.ibm.com/mx-es/topics/datamart>

Universidad Autónoma del Estado de Hidalgo. (n.d.). Retrieved from <https://www.uaeh.edu.mx/scige/boletin/icea/n9/r1.html>