

Evidencia de aprendizaje 2. Creación de una base de datos de Staging

Cristian Andrei Rendón Alcaraz

Kevin Saldarriaga Vélez

Santiago Agudelo Escobar

Ingeniería de Software y datos

Antonio Jesús Valderrama Jaramillo

Docente

Bases de Datos II-PREICA2502B010064

Curso

Institución Universitaria Digital de Antioquia

Antioquia

2025

Introducción

El siguiente documento propone una explicación de la base de datos sector de jardinería en staging, basada en el modelo relacional, el cual mostrara los datos relevantes y los que se trasladaran al bd staging, se diseñara, construirá y ejecutara las tablas, las consultas, registros que se piden en la actividad, permitiendo la validación de los datos u que estos puedan permanecer almacenados de manera correcta y optima, al final se elabora un backup de ambas bases de datos, la antigua y la nueva creada por staging.

Objetivos

Los principales objetivos dentro la elaboración de la base de datos staging, daremos a conocer e identificar los patrones comerciales los cuales se presentan en los datos operativos, como las ventas y los clientes recurrentes; elaborar un repositorio analítico el cual pueda centralizar la información guardada y actual de las ventas, clientes, productos.

También se busca facilitar la correcta toma de decisiones, a través de representaciones dinámicas del rendimiento y rentabilidad por región, la eficiencia de los vendedores y desempeño de los productos, finalizando con todo esto, se tomará la migración de datos buscando la optimización y posibilitando evitar la perdida o daño de datos, preservando la integridad de estos.

Planteamiento del problema

La base de datos actual presenta ciertas limitaciones, como lo son:

- Fragmentación analítica, ya que los datos claves están dispersos en las diferentes tablas normalizadas, forzando a hacer o realizar múltiple JOINS para consultas simples.
- Ausencia de historización, ya que algunas tablas solo registran el estado actual, lo cual hace imposible el seguimiento de los cambios, de un estado a otro.
- Dependencia de datos estáticos, hay varios datos los cuales son obsoletos, antiguos lo cual algunas veces genera reportes o datos inexactos.

Análisis del problema

La compañía Jardinería respecto a su base de datos a pesar de presentar una estructura eficaz para las operaciones, tiene algunas limitaciones con respecto a la complejidad de consultas, con sus múltiples tablas, retrasando acciones de ajustes o peticiones enfocadas con precisión en datos específicos importantes

- Datos desestructurados
- Limitaciones en cuanto al aprovechamiento de la información.
- Impedimento de identificación de patrones en cuanto a cruce de datos

- Impedimento para generar reportes automáticos en cuanto a la consolidación de datos importantes

Propuestas de solución

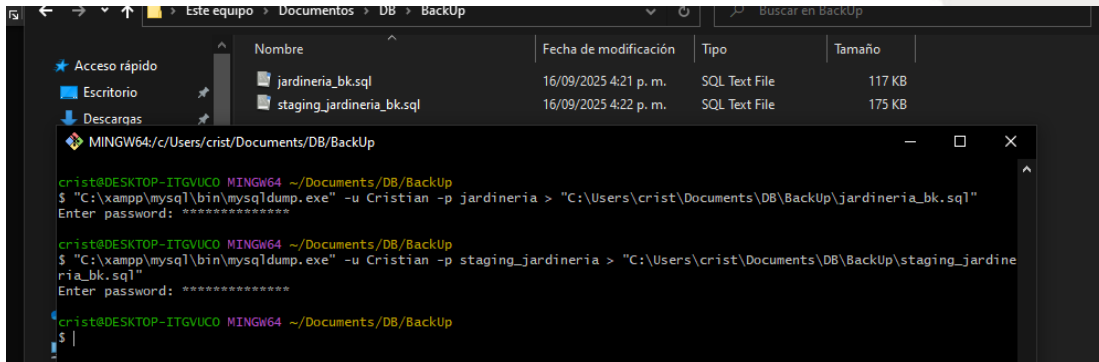
Se realizó un análisis exhaustivo de la base de datos origen jardineria, identificando todas sus tablas como relevantes para el proceso de integración de datos: oficina, empleado, cliente, producto, categoría_producto, pedido, detalle_pedido y pago. A partir de este análisis, se diseñó y construyó una nueva base de datos llamada staging_jardineria, completamente independiente del sistema, con el fin de crear un área de preparación aislada. La estructura de cada tabla en esta base de datos Staging fue creada para replicar fielmente las tablas origen, incorporando además metadatos clave como origen (para identificar la fuente de los datos), fecha_carga (marca temporal automática) y hash_data (un valor único generado mediante MD5 para detectar duplicados o cambios en futuras cargas). Este diseño permite no solo la migración segura de los datos, sino también su auditoría, validación y control. Posteriormente, se ejecutaron consultas SQL del tipo INSERT INTO ... SELECT ... para trasladar los registros desde las tablas de la base de datos origen hacia sus equivalentes en el área de staging, asegurando que todos los datos fueran transferidos correctamente. El proceso se llevó a cabo paso a paso para facilitar la detección de errores, y finalizó con una validación exhaustiva mediante conteos de registros que confirmaron la integridad y completitud de la carga. Finalmente, se generaron respaldos (backups) de ambas bases de datos utilizando la herramienta mysqldump, garantizando así la disponibilidad y recuperación de la información, culminando así un proceso completo y robusto de preparación de datos.

Conclusión

La puesta en marcha de una base de datos staging para el sistema de jardinería es un hito crucial en la modernización de su infraestructura de datos. Este proyecto no solo satisface las metas técnicas inmediatas, sino que también sienta las bases para un ecosistema analítico sólido y escalable. Con esto se dan a conocer los principales logros y sus repercusiones estratégicas:

- Se logra la validación al realizar la replicación precisa de tablas y relaciones garantiza que el staging opere como un espejo fiable de la base operacional, manteniendo las reglas del negocio y las restricciones originales.
- Al momento de la migración de los datos y registros, se logra evidenciar un correcto funcionamiento dando a concluir que hay riesgos de pérdida o corrupción de datos durante este, antes el staging facilita los procesos posteriores mejorando o enriqueciendo los datos.
- En cuanto a su arquitectura, se orienta a la escalabilidad, se permite la actualización independiente, se optimiza las cargas y posterior búsqueda.

Anexos



Jardinería_backup:

<https://drive.google.com/file/d/1bCfMJzkOInvTIAZamfMXVHjY9gKIqgkR/view?usp=sharing>

Staging_jardinería_backup:

https://drive.google.com/file/d/1hOwHxVhEmk66SF12cz3ivb_wvrfKQfZo/view?usp=sharing

Scripts con las consultas:

<https://drive.google.com/drive/folders/1dIQfkdo4PNOUkhtQ4rn6HRFUMhficM9X?usp=sharing>

Referencias

Kevininformático. (2017, 26 abril). Proceso ETL con SQL Server Integration Services |

Transferencia de Datos [Vídeo]. YouTube.

<https://www.youtube.com/watch?v=vSXplK9h8u8>

SAP Help Portal | SAP Online Help. (s. f.).

https://help.sap.com/docs/SAP_S4HANA_CLOUD/d5699934e7004d048c4801b552f3b013/47082e0d2baf40f5854f429dea179706.html?locale=es-ES

Tellado, F. (2025, 1 septiembre). ¿Cuándo usar staging, clonación o backup? Guía para elegir la mejor opción. SiteGround. <https://www.siteground.es/blog/cuando-usar-staging-clonacion-backup-wordpress/>