

EVIDENCIA DE APRENDIZAJE 2

CREACIÓN DE UNA BASE DE DATOS DE STAGING

INTEGRANTES

JHON JAIRO FUENTES TURIZO

GRUPO PREICA2502B010064

DOCENTE:

ANTONIO JESUS VALDERRAMA

INSTITUCIÓN UNIVERSIDAD IU DIGITAL DE ANTIOQUIA

INGENERIA DE SOFTWARE Y DATOS

2025

Tabla de contenido

EVIDENCIA DE APRENDIZAJE 2.....	1
Introducción	3
Objetivos	4
Planteamiento del problema	5
Análisis del problema.....	6
Propuesto de la solución	7
Descripción del análisis realizado a los datos Jardinería y cómo estos se trasladaron a la base de datos Staging	8
Estructura de la base de datos Staging.....	9
Consultas sql para copiar los registros de la base de datos jardineria a Staging.....	10
Backups de las base de datos	11
Conclusión:	12

Introducción

La presente actividad tiene como objetivo analizar la estructura y los datos de la base de datos Jardinería, con el fin de identificar la información más relevante para su traslado a una base de datos Staging. Este proceso es fundamental para preparar los datos antes de su integración en sistemas de análisis o almacenamiento definitivo, permitiendo una mejor organización, limpieza y transformación de la información. A través de la revisión de las tablas y sus relaciones, se busca seleccionar aquellos datos que aporten valor al proceso de toma de decisiones y al desarrollo de futuros proyectos académicos o profesionales relacionados con la gestión de bases de datos.

Objetivos

Objetivo general:

Analizar y seleccionar los datos relevantes de la base de datos Jardinería para su correcta migración y preparación en una base de datos Staging.

Objetivos específicos:

Identificar las tablas y campos esenciales de la base de datos Jardinería que aportan valor al proceso de análisis y toma de decisiones.

Proponer la estructura y las consultas SQL necesarias para crear una base de datos Staging que facilite la limpieza, transformación y posterior uso de los datos.

Planteamiento del problema

En el proceso de gestión y análisis de datos, es común que la información almacenada en las bases de datos operativas no se encuentre organizada ni depurada para su uso en sistemas de análisis o inteligencia de negocios. La base de datos Jardinería contiene una gran cantidad de datos distribuidos en diferentes tablas, lo que dificulta su integración y procesamiento eficiente. Por ello, surge la necesidad de identificar y seleccionar los datos más relevantes, así como de diseñar una base de datos Staging que permita preparar, limpiar y transformar la información antes de su migración al sistema de destino. Este planteamiento busca resolver los retos asociados a la calidad, organización y accesibilidad de los datos para facilitar su análisis y toma de decisiones.

Análisis del problema

El principal desafío radica en que los datos de la base Jardinería están distribuidos en múltiples tablas, cada una con diferentes tipos de información y relaciones. Esto puede generar inconsistencias, duplicidad y dificultades para acceder a los datos de manera eficiente. Además, no todos los datos almacenados son necesarios para el análisis, lo que obliga a realizar una selección cuidadosa de la información relevante. La falta de un proceso de limpieza y transformación previo puede afectar la calidad de los resultados obtenidos en etapas posteriores. Por ello, es fundamental analizar la estructura de la base de datos, identificar los datos clave y definir una estrategia para su migración y preparación en la base de datos Staging, asegurando así que la información esté lista para su uso en procesos de análisis y toma de decisiones.

Propuesto de la solución

Para abordar el problema identificado, se realizó un análisis detallado de la base de datos Jardinería, revisando la estructura y el contenido de sus tablas principales: clientes, empleados, productos, categorías, pedidos, detalles de pedidos y pagos. El objetivo fue identificar los datos esenciales para el proceso de análisis y toma de decisiones, descartando información redundante o poco relevante. A partir de este análisis, se seleccionaron las tablas y campos más importantes y se diseñó la estructura de una base de datos Staging. Esta base intermedia permite almacenar los datos de manera temporal, facilitando su limpieza, transformación y posterior migración al sistema de destino. El proceso incluyó la creación de scripts SQL para generar las tablas de Staging, asegurando que la información esté organizada y lista para su uso.

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

Se realizó un análisis de la estructura y contenido de la base de datos Jardinería, identificando un total de 8 tablas principales que almacenan información crítica del negocio:

oficina: Contiene datos de ubicación y contacto de las oficinas de la empresa.

empleado: Información del personal, incluyendo jerarquía organizacional y datos de contacto.

cliente: Datos completos de los clientes, información de contacto y límites de crédito.

Categoria_producto: Clasificación de productos por categorías.

producto: Inventario detallado con precios, proveedores y stock disponible.

pedido: Registro de órdenes realizadas por los clientes con fechas y estados.

detalle_pedido: Especificaciones de cada producto dentro de los pedidos.

pago: Historial de transacciones y métodos de pago utilizados.

Criterios de selección para la base de datos Staging

Durante el análisis se determinó que no todas las tablas eran igualmente relevantes para procesos de análisis y transformación de datos. Se aplicaron los siguientes criterios de selección:

- **Relevancia para análisis de negocio:** Tablas que aportan valor para la toma de decisiones
- **Frecuencia de consulta:** Datos utilizados regularmente en reportes y análisis
- **Calidad de datos:** Información completa y consistente

Estructura de la base de datos Staging

1. **stg_cliente**

Almacena los datos de los clientes relevantes para el análisis.

2. **stg_empleado**

Contiene la información de los empleados, útil para identificar representantes de ventas y responsables.

3. **stg_producto**

Guarda los datos de los productos, incluyendo categoría y precios.

4. **stg_categoria_producto**

Incluye las categorías de productos para clasificaciones posteriores.

5. **stg_pedido**

Registra los pedidos realizados por los clientes.

6. **stg_detalle_pedido**

Almacena el detalle de cada producto en los pedidos.

7. **stg_pago**

Contiene la información de los pagos realizados por los clientes.

Consultas sql para copiar los registros de la base de datos jardineria a Staging

-- Copiar clientes

```
INSERT INTO jardineria_staging.stg_cliente  
SELECT * FROM jardineria.cliente;
```

-- Copiar empleados

```
INSERT INTO jardineria_staging.stg_empleado  
SELECT * FROM jardineria.empleado;
```

-- Copiar productos

```
INSERT INTO jardineria_staging.stg_producto  
SELECT * FROM jardineria.producto;
```

-- Copiar categorías de productos

```
INSERT INTO jardineria_staging.stg_categoria_producto  
SELECT * FROM jardineria.Categoria_producto;
```

-- Copiar pedidos

```
INSERT INTO jardineria_staging.stg_pedido  
SELECT * FROM jardineria.pedido;
```

-- Copiar detalles de pedidos

```
INSERT INTO jardineria_staging.stg_detalle_pedido  
SELECT * FROM jardineria.detalle_pedido;
```

-- Copiar pagos

```
INSERT INTO jardineria_staging.stg_pago  
SELECT * FROM jardineria.pago;
```

Backups de las base de datos

Enlaces:

Base de datos jardineria:

<https://drive.google.com/file/d/1dXURX1673ma6y8T4cSWI5PKtoNZLJRyc/view?usp=sharing>

Base de datos jardineria_Stanging:

https://drive.google.com/file/d/1C-V8LTsUlj_54FQ4jHAGYWyz6GS6vm16/view?usp=sharing

Script de creación y consultas sql de la base de datos Stanging:

<https://drive.google.com/file/d/1zkACPp2d-yjOfTo53GDkq7wghClapQwM/view?usp=sharing>

Conclusión:

Durante este trabajo pude comprobar que crear una base de datos Staging para Jardinería fue realmente útil para organizar mejor los datos antes de analizarlos o pasarlos a otros sistemas. Al revisar cuidadosamente cómo estaba estructurada la base original y escribir las consultas SQL necesarias, logré construir un espacio intermedio donde es mucho más fácil limpiar, transformar y verificar la información. Lo que más me llamó la atención es que este enfoque no solo hace que el manejo de datos sea más eficiente, sino que también deja todo preparado para trabajos futuros relacionados con análisis de datos y reportes empresariales. Definitivamente es una metodología que aplicaría en otros proyectos similares.

La implementación exitosa de las siete tablas de staging (stg_cliente, stg_empleado, stg_producto, stg_categoria_producto, stg_pedido, stg_detalle_pedido y stg_pago) permitió trasladar de manera íntegra los datos más relevantes de la base original, manteniendo la consistencia y calidad de la información. Las consultas INSERT INTO SELECT desarrolladas garantizaron una migración completa y exacta de los registros.