



**Nombre:** Elkin Mauricio Bautista Rubiano

**Clase:** PREICA2502B010064

**Docente:** Antonio Jesús Valderrama Jaramillo

**Materia:** Bases de datos II

## Introducción:

En el campo de la ingeniería de datos, la gestión adecuada de la información es crucial para garantizar la integridad y disponibilidad de los sistemas. Un paso fundamental en los flujos de trabajo de extracción, transformación y carga (ETL) es la implementación de un área de preparación o *staging*. Este entorno intermedio permite aislar los datos operativos de los procesos de análisis y transformación, minimizando los riesgos sobre la base de datos de producción.

Este informe detalla el proceso técnico seguido para la creación de una base de datos de *staging* a partir del esquema existente de "Jardinería". Se describe el análisis inicial, el diseño de la solución, la implementación mediante scripts SQL y la validación de los resultados, reflejando los desafíos y aprendizajes encontrados durante el proceso.

## Objetivos:

Implementar una base de datos de staging funcional para la base de datos jardineria, asegurando una transferencia de datos íntegra y validada desde el entorno de origen.

### Objetivos Específicos

- Diseñar la estructura de tablas para la base de datos de staging basándose en el esquema de la base de datos jardineria.
- Construir y documentar las consultas SQL necesarias para migrar los registros desde la base de datos original a la de staging.
- Ejecutar el proceso de migración y validar que los datos se hayan almacenado correctamente en el nuevo entorno.
- Generar copias de seguridad (backups) tanto de la base de datos original como de la de staging para garantizar la recuperabilidad.

## Planteamiento del Problema

La base de datos jardineria contiene información operativa valiosa que es esencial para el funcionamiento diario del negocio. Cualquier proceso de análisis, limpieza o transformación de datos que se ejecute directamente sobre ella podría comprometer su integridad, afectar su rendimiento o, en el peor de los casos, provocar una pérdida de información.

Por ello, surge la necesidad de crear un entorno intermedio y aislado, conocido como base de datos de staging. Este espacio permitirá manipular y preparar los datos para futuros análisis o para su carga en un data warehouse, sin impactar directamente el sistema de producción.

## Análisis del Problema

El análisis se centró en comprender la estructura de la base de datos jardineria original. Se revisaron las tablas existentes, sus columnas, tipos de datos y las relaciones de clave foránea que las conectan. Aunque mi experiencia previa con sistemas ERP como Odoo me ha dado una base sobre cómo se estructuran los datos a nivel de aplicación, el manejo directo de scripts SQL para este tipo de tareas de migración fue un reto interesante.

Se determinó que, para esta fase inicial del proyecto, la estrategia más segura y directa sería replicar la estructura existente en un nuevo esquema de staging. Este enfoque 1:1 asegura que no se pierda información en la transferencia inicial y proporciona una base sólida y conocida sobre la cual se podrán aplicar transformaciones en el futuro.

## Propuesta de la solución:

La solución se implementó siguiendo una serie de pasos lógicos, desde la creación de la estructura hasta la validación final de los datos.

## Correcciones a la Entrega 1

Durante la fase de construcción de los scripts de migración, me encontré con un desafío técnico. Inicialmente, intenté utilizar consultas del tipo `INSERT INTO ... SELECT *` para agilizar la transferencia de datos. Sin embargo, esto generó un error

SQL Error [1136] [21S01]: Column count doesn't match value count at row 1.

Tras investigar el error, comprendí que esta práctica, aunque rápida, es poco robusta. La solución fue modificar las consultas para especificar explícitamente las columnas tanto en la cláusula `INSERT INTO` como en el `SELECT`. Por ejemplo:

Este ajuste no solo solucionó el error, sino que también hizo los scripts más mantenibles y menos propensos a fallar si la estructura de las tablas cambia en el futuro. Fue un aprendizaje práctico muy valioso.

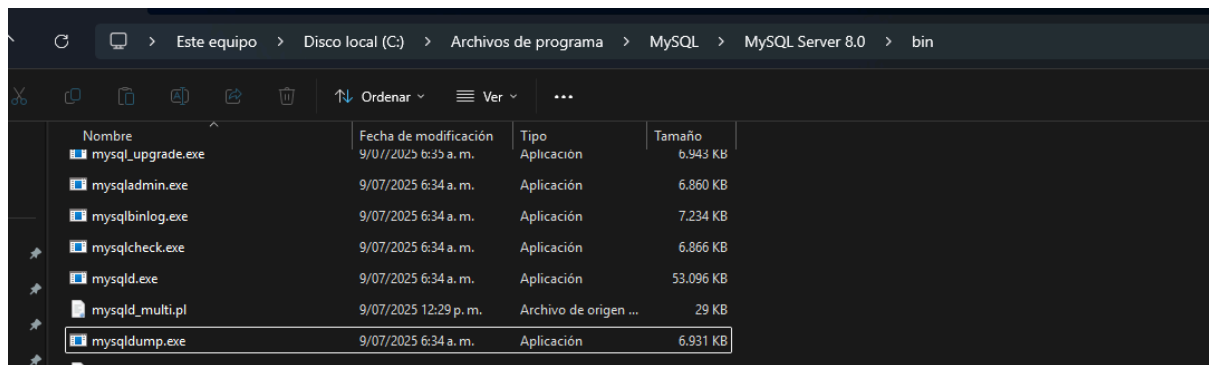
## Creación de backups

Para crear los backups, tanto de la base de datos común y corriente llamada jardineria, como la de jardineria\_staging, tuve que navegar mediante cmd mediante los comandos:

```
CD ../../
```

CD Program Files\MySQL\MySQL Server 8.0\bin

Dentro de la carpeta bin, pude encontrar el archivo mysqldump.exe



Luego, en CMD, ejecuté la siguiente línea de comandos:

```
mysqldump -u [usuario] -p nombreBaseDeDatosActual > nombreArchivoBackup.sql
```

Esta línea recibe varios argumentos, al lado de -u el usuario, que en MySQL casi siempre será ROOT, en -p simbolizamos la contraseña, seguido de ese -p indicamos el nombre de nuestra base de datos y después colocamos un > seguido de el nombre que llevará el archivo de nuestro backup; después de esto el motor MySQL nos pide nuestra contraseña.

Ejemplo diligenciado para ambos casos:

Base de datos Jardineria:

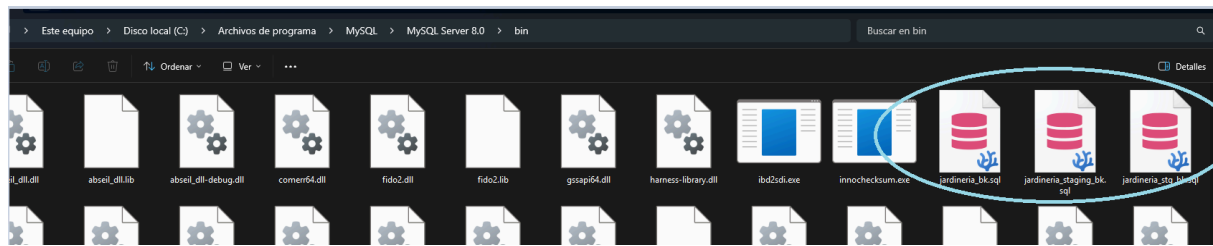
```
mysqldump -u root -p jarineria > jardineria_bk.sql
```

Base de datos Jardineria\_staging\_bk.sql:

```
mysqldump -u root -p jardineria_staging > jardineria_staging_bk.sql
```

Resultado:

Ahora podremos encontrar los archivos dentro del mismo directorio en el que encontramos el archivo mysqldump.exe



Bibliografía:

<https://www.youtube.com/watch?v=oKmBvO8rUdo>