

Autor : Angel Teodoro Jaramillo Sulca

Informe – DWH TechStore

Se implementó un Data Warehouse de tres capas (Staging, ODS, BDS) para TechStore, con un modelo dimensional en estrella y un paquete SSIS que automatiza la carga. Las dimensiones se cargaron con lógica SCD2, la DimTiempo se generó correctamente y la FactVentas se alimenta mediante joins a las dimensiones. La principal dificultad se presentó en la **capa ODS**, al transformar datos desde Staging.

Diseño y transformaciones

- **Staging:** Datos crudos, sin validación.
- **ODS:** Normalización y limpieza. Aquí se aplicaron conversiones de tipos, deduplicación y cálculos.
- **BDS:** Esquema estrella con DimProducto, DimCliente, DimTienda, DimTiempo y FactVentas.

Transformaciones clave:

- **ODS.Ventas:** Conversión de **FechaVenta** (VARCHAR → DATE), cálculo de **MontoTotal**.
- **ODS.Clientes:** Email en minúsculas.
- **ODS.Tiendas:** Normalización de campo **Tipo**.
- **ODS.Productos:** Control de actualizaciones.

Dificultades encontradas

- **ODS (principal reto):**
 - Al crear columnas derivadas en SSIS, los tipos y longitudes no coincidían con lo que esperaba la tabla ODS.
 - Esto generó errores de carga al pasar de Staging a ODS.
 - Se resolvió ajustando conversiones explícitas (ej. **DT_WSTR** → **DT_STR**, **VARCHAR** → **DATE**) y alineando nombres/longitudes con la definición de ODS.
- **FactVentas:** Inicialmente se intentó usar **MERGE** con **VentaID**, pero al no existir en la tabla de hechos, se corrigió con **INSERT** controlado y joins a dimensiones.

Optimización aplicada

- Índices en claves de negocio + EsActual en dimensiones para acelerar lookups.
- Índices en foreign keys de FactVentas.
- Uso de transacciones y TRY/CATCH en cargas SCD2.

Automatización SSIS

- **Control Flow:** Secuencia de limpieza de Staging, carga a Staging, procesamiento ODS, carga de dimensiones y hechos en BDS.
- **Data Flow FactVentas:** Source ODS.Ventas → Lookups a dimensiones → cálculo de MontoTotal → Destination FactVentas.

Conclusión

El DWH cumple con el objetivo de integrar datos de ventas, productos, clientes y tiendas en un modelo dimensional optimizado para análisis. La mayor dificultad se concentró en la **capa ODS**, donde fue necesario alinear cuidadosamente las columnas derivadas con los tipos y longitudes esperados en las tablas destino. Una vez resuelto, el flujo completo se ejecutó de manera estable.

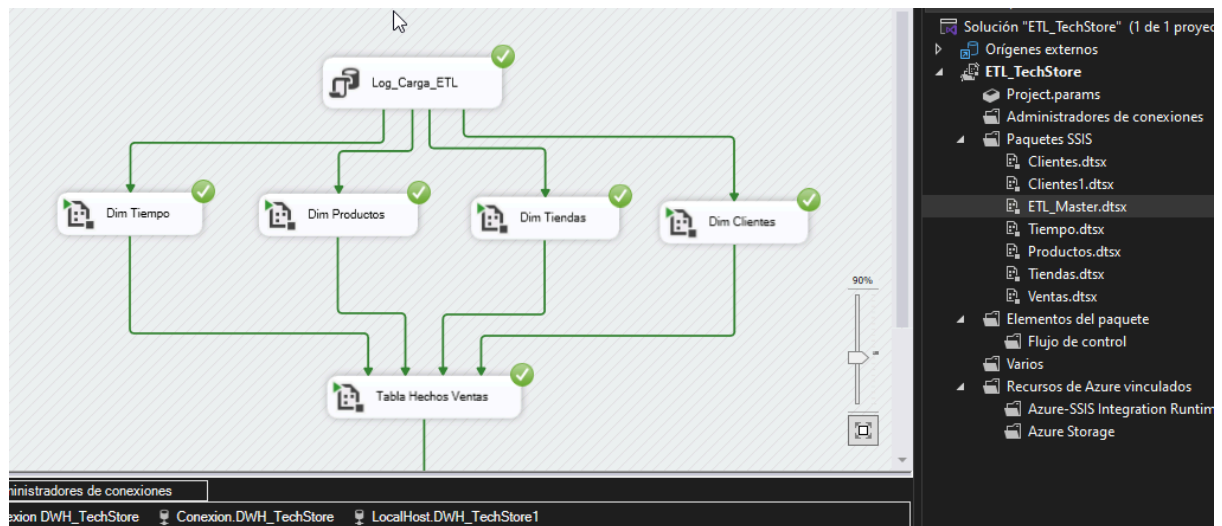


Image - Producto

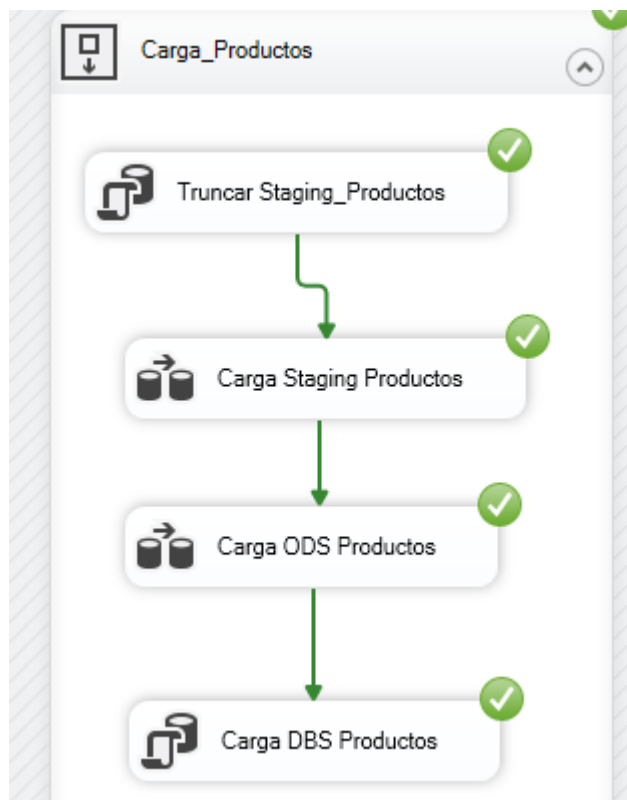


Image - Clientes

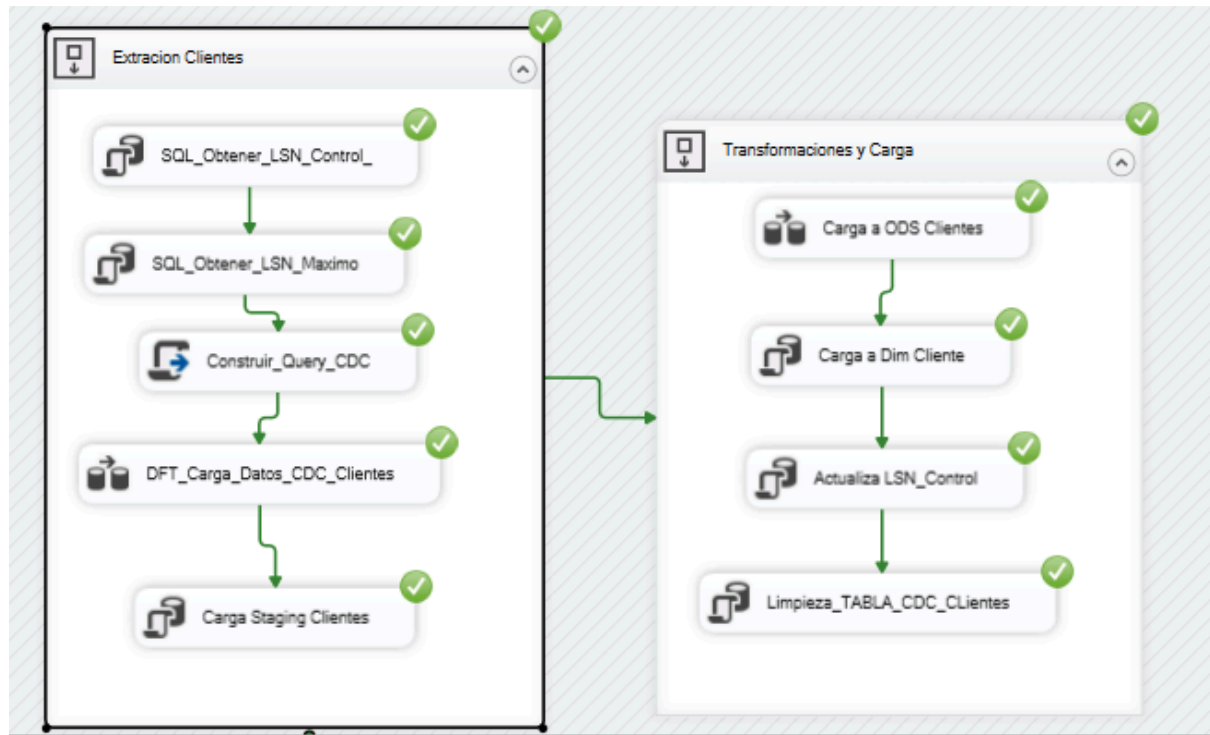


Image - Cliente

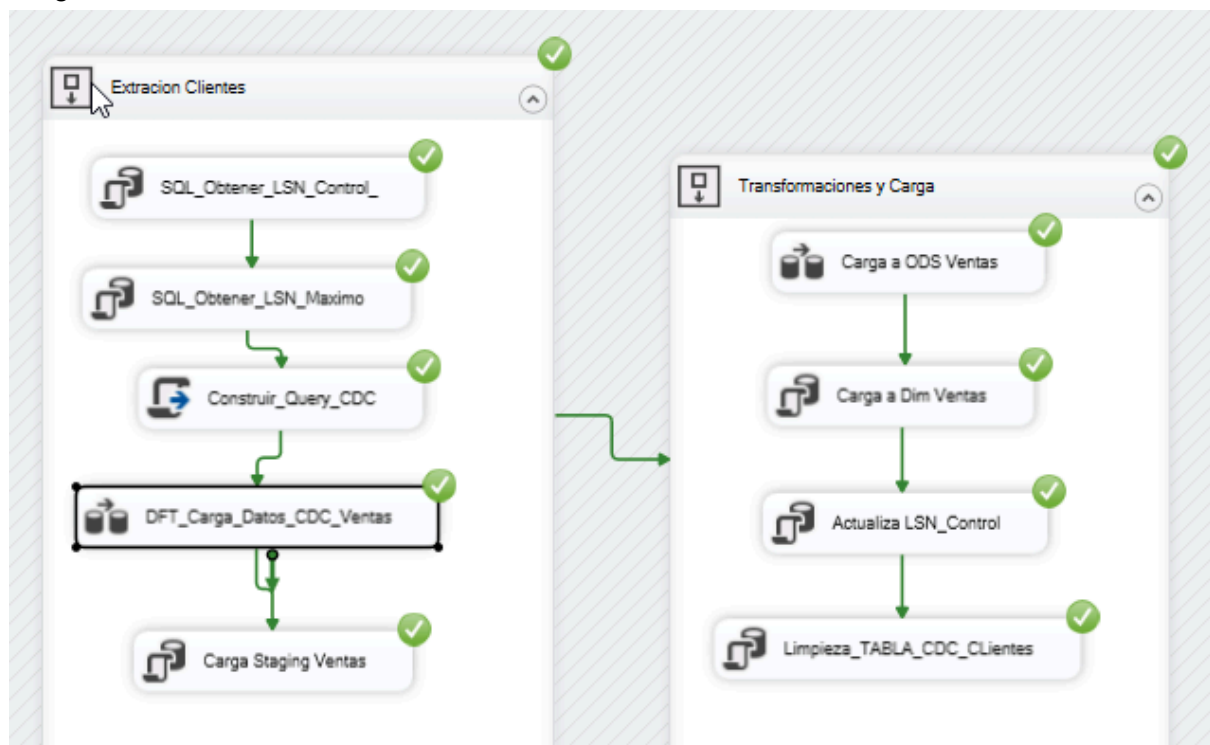


Image - Conteo de Datos para cada Tabla

```
121 |
122 | SELECT 'Productos' AS Tabla, COUNT(*) AS Total FROM Fuente.Productos
123 | UNION ALL
124 | SELECT 'Clientes', COUNT(*) FROM Fuente.Clientes
125 | UNION ALL
126 | SELECT 'Tiendas', COUNT(*) FROM Fuente.Tiendas
127 | UNION ALL
128 | SELECT 'Ventas', COUNT(*) FROM Fuente.Ventas;
```

00 % 11 0

Resultados Mensajes

| | Tabla | Total |
|---|-----------|-------|
| 1 | Productos | 20 |
| 2 | Clientes | 30 |
| 3 | Tiendas | 10 |
| 4 | Ventas | 100 |

Image Ventas por Mes

```
129 |
130 | SELECT
131 |     YEAR(CAST(FechaVenta AS DATE)) AS Anio,
132 |     MONTH(CAST(FechaVenta AS DATE)) AS Mes,
133 |     COUNT(*) AS TotalVentas
134 | FROM Fuente.Ventas
135 | GROUP BY YEAR(CAST(FechaVenta AS DATE)), MONTH(CAST(FechaVenta AS DATE))
136 | ORDER BY Anio, Mes;
```

100 % 16 0

Resultados Mensajes

| | Anio | Mes | TotalVentas |
|----|------|-----|-------------|
| 1 | 2024 | 1 | 10 |
| 2 | 2024 | 2 | 10 |
| 3 | 2024 | 3 | 10 |
| 4 | 2024 | 4 | 10 |
| 5 | 2024 | 5 | 10 |
| 6 | 2024 | 6 | 10 |
| 7 | 2024 | 7 | 10 |
| 8 | 2024 | 8 | 10 |
| 9 | 2024 | 9 | 10 |
| 10 | 2024 | 10 | 10 |