

Documentación ETL – Dim_Localizacion

1. Objetivo

Construir la dimensión **Dim_Localizacion** que consolide información geográfica y de segmento (City, State, Country, PostalCode, Region, Segment), asegurando unicidad, consistencia y trazabilidad para el análisis de ventas por ubicación.

2. Pasos realizados

- **Carga inicial:** Se partió del archivo localizacion.csv, promoviendo encabezados y tipando columnas (City, Segment, Country, State, PostalCode, Region).
- **Corrección de nombres de columnas:** Se eliminaron espacios en los nombres (Postal Code → PostalCode) para mantener consistencia con el resto del modelo.
- **Normalización de atributos:**
 - City, State, Country, Region, Segment → capitalizados con Text.Proper.
 - PostalCode → tipado como texto o entero según necesidad, asegurando trazabilidad.
- **Validaciones de calidad:**
 - **Duplicados:** se detectaron ciudades repetidas con distintos segmentos.
 - **Nulos críticos:** se identificaron registros sin City, Country o Region.
 - **Consistencia:** se verificó que una misma ciudad no esté asignada a más de un país o región.
- **Consolidación automática:** Se agruparon registros por City + State + Country + PostalCode, tomando el primer valor de Segment y Region como referencia.
- **Tabla de auditoría (AuditDimLocalizacion):** Se construyó un resumen con tres métricas:
 - **Duplicados detectados** → cuántas localizaciones aparecían más de una vez.
 - **Nulos críticos** → cuántos registros tenían problemas en campos clave.
 - **Localizaciones consolidadas** → cantidad final de filas únicas en la dimensión.

3. Resultados

- **Dim_Localizacion** queda con una fila única por ciudad/estado/país/código postal.
- Segmento y región se consolidan automáticamente.
- La tabla AuditDimLocalizacion documenta duplicados, nulos y cantidad final de localizaciones.
- La dimensión está lista para relacionarse con **Fact_Sales** y **Dim_Customers**, aportando trazabilidad geográfica al modelo estrella.