

Manual de Diccionario de Datos

Activos de Información: TechLogistics S.A.

Curso: Fundamentos en Ciencia de Datos
Docente: Jorge Iván Padilla-Buriticá
Universidad EAFIT - Periodo 2026-1

Importancia del Diccionario

En la consultoría senior, el diccionario de datos no es solo una lista de columnas; es el mapa de los procesos de negocio. Entender el **porqué** de una captura permite identificar si una anomalía es un error técnico o un síntoma de un fallo operativo.

1 Dataset 1: Inventario Central (inventario_central_v2.csv)

Este archivo representa el stock teórico reportado por el sistema ERP al inicio del trimestre.

Variable	Descripción	Propósito / Razón de Captura
SKU_ID	Identificador único de producto.	Trazabilidad: Llave primaria para cruzar con ventas.
Categoria	Segmento del mercado.	Estrategia: Agrupar productos para análisis de márgenes.
Stock_Actual	Unidades físicas en bodega.	Operación: Evitar quiebres de stock. <i>Alerta: Valores negativos.</i>
Costo_Unitario_USD	Precio de compra al proveedor.	Contabilidad: Base para calcular el margen bruto.
Punto_Reorden	Nivel crítico de inventario.	Logística: Disparador automático de compras.
Lead_Time_Dias	Tiempo de entrega del proveedor.	Planeación: Tiempo esperado de reposición. <i>Dato ruidoso.</i>
Bodega_Origen	Nodo de despacho.	Geografía: Identificar centros de costos regionales.
Ultima_Revision	Fecha del último conteo físico.	Confianza: Indica qué tan actualizado está el dato.

2 Dataset 2: Transacciones Logística (transacciones_logistica_v2.csv)

Contiene el flujo de salida de mercancía y la ejecución de la promesa de entrega.

Variable	Descripción	Propósito / Razón de Captura
Transaccion_ID	ID único de orden de venta.	Control: Registro individual de cada operación.
SKU_ID	ID del producto vendido.	Integridad: Cruce de salida de stock. <i>Alerta: IDs inexistentes.</i>
Fecha_Venta	Momento de la transacción.	Temporalidad: Análisis de estacionalidad y demanda.
Cantidad_Vendida	Volumen de unidades.	Ventas: Cálculo de ingresos y agotamiento de stock.
Precio_Venta_Final	Precio cobrado al cliente.	Ingresos: Valor transaccional real post-descuentos.
Costo_Envio	Gasto de transporte.	Rentabilidad: Impacta directamente el margen neto.
Tiempo_Entrega	Días desde pedido a entrega.	Servicio: KPI principal de logística. <i>Alerta: Outliers.</i>
Estado_Envio	Situación actual del pedido.	Visibilidad: Seguimiento de la última milla.
Ciudad_Destino	Ubicación del cliente.	Mercadeo: Análisis de penetración regional.

3 Dataset 3: Feedback Clientes (feedback_clientes_v2.csv)

Datos cualitativos y cuantitativos de la experiencia post-venta.

Variable	Descripción	Propósito / Razón de Captura
Feedback_ID	Registro del comentario.	Auditoría: Control de atención al cliente.
Transaccion_ID	Referencia a la venta.	Relación: Atribuir la opinión a un producto/bodega.
Rating_Producto	Calificación 1-5.	Calidad: Identificar productos defectuosos.
Rating_Logistica	Calificación del envío.	Desempeño: Evaluar a los transportistas.
Satisfaccion_NPS	Índice de recomendación.	Lealtad: Predecir la retención de clientes a largo plazo.
Ticket_Soporte	Si hubo reclamo formal (Sí/No).	Carga Operativa: Costo oculto de una mala venta.

Nota para el Consultor

TechLogistics S.A. captura estos datos para maximizar la rentabilidad. Sin embargo, si un **SKU_ID** no existe en inventario pero tiene ventas, estamos ante una **Venta Fantasma** que el consultor debe diagnosticar: ¿Falla de catálogo o fraude?