UNIVERSIDAD DE EL SALVADOR FACULTAD DE INGENIERÍA Y ARQUITECTURA ESCUELA DE INGENIERÍA DE SISTEMAS INFORMÁTICOS.



CURSO DE ESPECIALIZACIÓN INGENIERÍA DE DATOS Ing. Rene Fabricio Quintanilla Gomez

Primera fase: Diseño de un modelo multidimensional para soportar un proceso de negocio seleccionado

Tema: Construcción de data warehouse para análisis de inventarios para las sociedades de grupo probe.

Grupo: 01

Carnet	Apellidos	Nombres
BD03006	Beltrán Domínguez	José Eduardo
EQ17001	Escobar Quintanilla	Francisco Josué
PG14048	Palacios González	Gonzalo de Jesús
RP18045	Reyes Perdido	German Oswaldo

Introducción

Este documento tiene como objetivo presentar el diseño de un modelo multidimensional (Data Warehouse) para soportar el análisis de los procesos transaccionales de inventario de la empresa Probe, utilizando la base de datos transaccional del sistema erp softland de la empresa. El objetivo principal es proporcionar una plataforma analítica que nos permita mostrar información de carácter gerencial bajo análisis multidimensional para tomar decisiones sobre la eficiencia de los ingresos, ventas, consumos , el rendimiento de los proveedores, los costos asociados y movimiento del inventario.

El diseño se basa en una arquitectura de esquema estrella (Star Schema), reconocida por su simplicidad para la comprensión del usuario final y su optimización para el rendimiento de consultas analíticas.

Lógica de Negocio y Objetivos del Análisis

La lógica de negocio propuesta busca resolver los desafíos derivados de la fragmentación de los sistemas operativos de Grupo Probe y consolidar, bajo un enfoque analítico, toda la información relacionada con la gestión de inventarios en sus distintas sociedades de negocio.

El Grupo maneja inventarios que cumplen una doble función: como insumos para servicios y como productos comerciales. Por ello, el modelo de negocio requiere un esquema de control que permita rastrear cada producto desde su Ingreso hasta su consumo o venta final.

Las transacciones se originan en sistemas distintos: RMS, OVST, Softland POS. Estas fuentes manejan sus catálogos, estructuras y reglas operativas similares, por lo cual la solución debe homologar y transformar la información para integrar bajo un modelo común.

El modelo de datos propuesto se estructura en torno a cuatro procesos principales que abarcan todo el ciclo de vida del inventario:

- Ingresos: productos adquiridos a proveedores que ingresan a bodega.
- Traslado: movimientos internos entre bodegas de una misma sociedad.
- **Consumos:** productos utilizados en la prestación de servicios para sala(vidals y sento) en caso de distribuidora se definiría como insumo para regalía o complemento a la venta.
- Ventas: productos vendidos al cliente final.

El objetivo principal de esta lógica de negocio es establecer un modelo analítico común que permita:

- Integrar las transacciones de inventario provenientes de múltiples sistemas operativos (RMS, OVST, Softland POS).
- Unificar la información bajo un modelo de datos estructurado y centralizado.
- Representar fielmente los procesos de Ingreso, traslado, consumo y venta de inventario.
- Brindar una base confiable para la trazabilidad de productos y el análisis avanzado de información.

Los objetivos de negocio que este modelo multidimensional trata de dar solución son:

Ingreso de articulos (ingresos de inventario)

Venta de artículos (salidas de inventario)

Consumo de artículos (salida de inventario)

Transferencias entre bodegas (movimiento entre inventarios)

Reglas de Negocio y Consideraciones Generales

Reglas Generales para Dimensiones

• DimFecha:

- Debe contener todas las granularidades temporales necesarias para el análisis (día, semana, mes, trimestre, año).
- Debe incluir atributos como día de la semana, nombre del mes, número de semana del año, feriados, etc., para facilitar el análisis temporal.
- Es una dimensión compartida por todas las tablas de hechos, asegurando consistencia temporal en los análisis.

DimArticulo:

- Debe contener información detallada sobre cada artículo (código, nombre, descripción, unidad de medida, categoría, subcategoría, marca, precio de venta, costo promedio).
- Debe manejar cambios en los atributos del artículo a lo largo del tiempo (ej. cambios de precio, descripciones). Considerar una Dimensión de Cambio Lento (SCD Tipo 2) si se requiere historial.
- Es una dimensión compartida por todas las tablas de hechos que involucran inventario, asegurando consistencia en la descripción del producto.

DimBodega:

- Debe contener información sobre cada ubicación de almacenamiento (código de bodega, nombre, dirección, tipo de bodega, capacidad).
- Es una dimensión compartida por todas las tablas de hechos que manejan movimientos de inventario, fundamental para el análisis de existencias y flujos.

• DimCliente:

- Debe contener información detallada del cliente (ID, nombre, tipo de cliente, dirección, teléfono, correo electrónico, segmentación, fecha de alta).
- Debe manejar cambios en los atributos del cliente (ej. cambio de dirección, segmento). Considerar una Dimensión de Cambio Lento (SCD Tipo 2).
- Exclusiva de FactVentas: Indica que solo las ventas están directamente asociadas a un cliente final.

DimProveedor:

- Debe contener información detallada del proveedor (ID, nombre, contacto, dirección, términos de pago).
- Debe manejar cambios en los atributos del proveedor. Considerar una Dimensión de Cambio Lento (SCD Tipo 2).
- Exclusiva de FactIngresos: Indica que solo las Ingresos están directamente asociadas a un proveedor.

DimLote:

- Debe contener atributos específicos del lote (número de lote, fecha de fabricación, fecha de caducidad, cantidad original del lote).
- Es crucial para trazabilidad y gestión de inventario, especialmente si se manejan productos perecederos o con número de serie.
- Compartida por FactVentas, FactTraspaso y FactConsumo, indicando la importancia de la trazabilidad de lotes en estos procesos.

. Reglas Específicas por Tabla de Hechos

FactVentas:

- Granularidad: Cada fila representa un detalle de línea de venta (un artículo vendido en una transacción específica).
- Métricas: Cantidad vendida, precio unitario de venta, valor total de venta, costo unitario, margen de beneficio, descuento aplicado.

Asociaciones:

- Cada venta debe tener una Fecha (DimFecha).
- Cada línea de venta debe referenciar un Artículo (DimArticulo).
- Cada venta debe originarse o afectar una Bodega (DimBodega).
- Cada venta debe estar asociada a un Cliente (DimCliente).
- Si aplica la trazabilidad, cada línea de venta debe estar asociada a un Lote (DimLote).
- Validación: Asegurar que los precios de venta son válidos, que las cantidades no son negativas.

• FactIngresos:

- Granularidad: Cada fila representa un detalle de línea de Ingreso (un artículo Ingreso de en una orden de Ingreso específica).
- Métricas: Cantidad Ingreso de, precio unitario de Ingreso, valor total de Ingreso, impuestos aplicados, costos de flete.

Asociaciones:

- Cada Ingreso debe tener una Fecha (DimFecha).
- Cada línea de Ingreso debe referenciar un **Artículo** (DimArticulo).
- Cada Ingreso debe tener una Bodega de destino (DimBodega).
- Cada Ingreso debe estar asociada a un Proveedor (DimProveedor).
- Validación: Asegurar que los precios de Ingreso son válidos, que las cantidades no son negativas.

FactTraspaso:

Granularidad: Cada fila representa un movimiento de un artículo entre dos bodegas.

Métricas: Cantidad traspasada.

Asociaciones:

- Cada traspaso debe tener una Fecha (DimFecha).
- Cada traspaso debe referenciar un **Artículo** (DimArticulo).

Validación: La cantidad traspasada no puede ser negativa. El artículo debe existir en la bodega origen en la cantidad especificada.

FactConsumo:

- Granularidad: Cada fila representa el consumo de un artículo de inventario por algún proceso interno (ej. producción, muestras, mermas).
- **Métricas:** Cantidad consumida, costo del consumo.

Asociaciones:

- Cada consumo debe tener una Fecha (DimFecha).
- Cada consumo debe referenciar un Artículo (DimArticulo).
- Cada consumo debe originarse de una Bodega (DimBodega).
- Si aplica la trazabilidad, cada consumo debe estar asociado a un Lote (DimLote).
- **Validación:** La cantidad consumida no puede ser negativa. El artículo debe existir en la bodega de origen en la cantidad especificada.

Diseño Detallado de Dimensiones y Tablas de Hechos

Basado en la Matriz de Bus, el modelo multidimensional consta de cuatro tablas de hechos y seis dimensiones compartidas que nos servirán para el análisis del modelo dimensional.

Bus Matriz

	DimFecha	DimArticulo	DimBodega	DimCliente	DimProveedor	DimLote
FactVentas	х	Х	х	х		х
FactIngresos	х	Х	х		х	х
FactTraspaso	х	Х	х			х
FactConsumo	х	Х	х			х

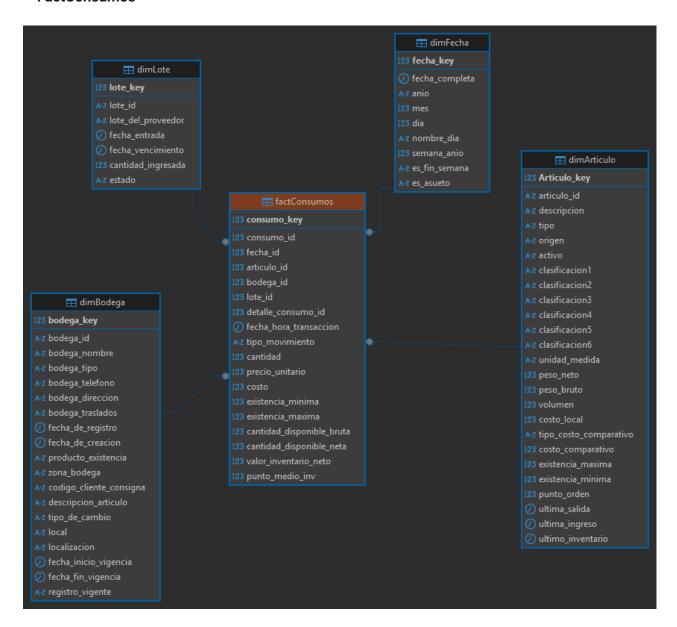
Consideraciones generales aplicada a las dimensiones y tablas de hecho

- Manejo de Errores ETL: La robustez del Data Warehouse dependerá de la calidad del proceso ETL para manejar errores de datos, transformaciones fallidas y asegurar la integridad referencial.
- Rendimiento: Para conjuntos de datos muy grandes, se pueden considerar técnicas como el particionamiento de tablas de hechos y la agregación precalculada (cubos OLAP) para mejorar el rendimiento de las consultas.
- Calidad de Datos: Dependerá mucho del ingreso de datos en el ERP de Softland, bajo la cual se realizará una limpieza de datos preliminar antes de la generación de ETLs.
- Extensión a Otros Procesos: La adopción de la Matriz de Bus (como la proporcionada)
 facilita la extensión futura del Data Warehouse a otros procesos de negocio (ej. Ventas,
 Recursos Humanos) reutilizando dimensiones existentes y añadiendo nuevas tablas de
 hechos. El data warehouse que se desarrolle se replicará por la empresa para todas las
 sociedades que posee

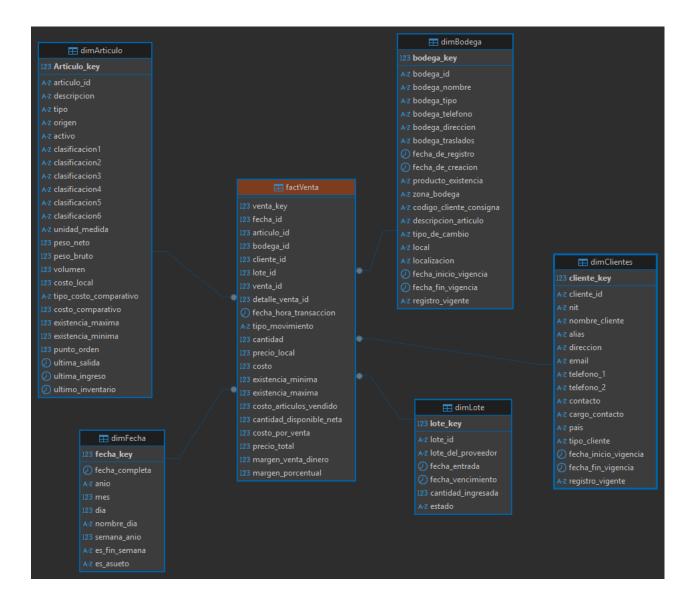
Las dimensiones utilizadas incluyen campos para el manejo de Slowly Changing Dimensions (SCD) Tipo 2 (como fecha_inicio_vigencia, fecha_fin_vigencia, registro_vigente), lo cual es una excelente práctica en Data Warehousing el cual utilizamos para mantener el historial de los cambios en los atributos de las dimensiones.

Diagrama del modelo

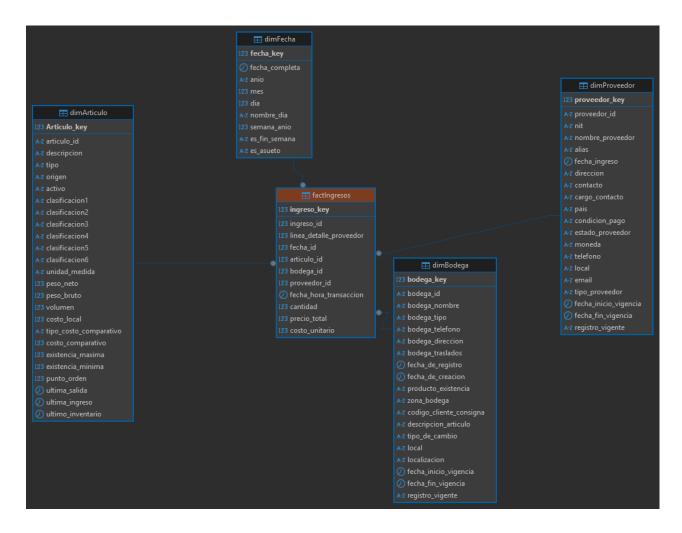
FactConsumos



FactVenta



FactIngresos



FactTraspaso



Diccionario de datos

TABLA: DimClientes

Campo	Tipo	Validación	Descripción
cliente_key	int	AutoincrementalLlave primaria	Llave surroga de la tabla dimClientes
cliente_id	varchar(20)	No nulo	Llave primaria del sistema transaccional
NIT	varchar(20)	No nulo	Numero de identificacion tributaria del cliente
nombre_cliente	varchar(150)	 No nulo 	Nombre de la persona jurídica o natural
alias	varchar(150)	No nulo	Nombre con el que es conocida la empresa
direccion	text	Default: N/A	Ubicación donde se encuentra el cliente
contacto	varchar(30)	No nulo	Nombre del contacto del cliente
cargo_contacto	varchar(30)	No nulo	Nombre del cargo del contacto del cliente
email	varchar(256)	 No nulo 	Correo electronico del cliente
telefono_1	varchar(50)	 Default: No posee numero de contacto 	Primera opcion de telefono de contacto
telefono_2	varchar(50)	Default: No posee numero de contacto	Segunda opcion de telefono de contacto

Pais	varchar(40)	 No nulo 	Nombre del país de origen del cliente
tipo_cliente	varchar(40)	No nulo	Nombre de la categoría del cliente asignada por la empresa.
fecha_inicio_vigencia	date	Default: fecha actual	Fecha en que entra en vigencia el registro
fecha_fin_vigencia	date	Acepta nulos	Fecha en que finaliza la vigencia del registro
registro_vigente	varchar(2)	Default: SI	Variable que indica si es un registro vigente

TABLA: DimFechas

Campo	Tipo	Validación	Descripción
fecha_key	int	AutoincrementalLlave primaria	Llave surroga de la tabla dimFechas
fecha_completa	date	No nulo	fecha completa en formato dia/mes/año
anio	varchar(4)	 No nulo 	Año en número
mes	varchar(2)	No nulo	Mes en número
dia	varchar(2)	 No nulo 	Dia en número
nombre_dia	varchar(50)	No nulo	Nombre del dia de la semana
semana_anio	int	No nulo	Indica la semana del año
es_fin_semana	varchar(2)	No nuloDefault: NO	Indica si es una fecha de fin de semana

€	es_asueto	varchar(2)	•	No nulo Default: NO	Indica si es una fecha feriada

TABLA: FactConsumos

Campo	Tipo	Validación	Descripción
consumo_key	int	AutoincrementalLlave primaria	Llave surroga de la tabla factConsumos
consumo_id	int	No nulo	Llave primaria del sistema transaccional
detalle_id	int	No nulo	Llave primaria del detalle del sistema transaccional
fecha_id	int	Llave foránea	Llave foránea de DimFecha
articulo_id	int	Llave foránea	Llave foránea de DimArticulo
bodega_id	int	Llave foránea	Llave foránea de DimBodega
lote_id	int	Llave foránea	Llave foránea de DimLote
fecha_hora_transaccion	datetime	No nulo	Fecha y hora de registro de transacción
tipo_movimiento	varchar(15)	Default: Consumo	Nombre del tipo de movimiento realizado
cantidad	decimal(28,13)	No nulo	Valor solicitado del artículo
precio_unitario	decimal(28,13)	No nulo	Valor monetario del artículo
costo	decimal(28,13)	No nulo	Costo del consumo calculado por cantidad * precio_unitario

existencia_minima	decimal(28,13)	No nuloDefault: 0	Existencia mínima del producto
existencia_maxima	decimal(28,13)	No nuloDefault: 0	Existencia máxima del producto
cantidad_disponible_bruta	decimal(28,13)	No nulo	Cantidad disponible antes del consumo
cantidad_disponible_neta	decimal(28,13)	No nulo	Cantidad disponible después del consumo
valor_inventario_neto	decimal(28,13)	 No nulo 	Valor monetario de la cantidad de artículos almacenados después del consumo
punto_medio_inv	decimal(28,13)	 No nulo 	Nivel de inventario promedio calculado, indica en qué momento queda desabastecido por consumo

TABLA: factIngresos

Campo	Tipo	Validación	Descripción
Ingresos_key	int	AutoincrementalLlave primaria	Llave surrogada de la tabla FactIngresos
Ingresos_id	int	No nulo	Identificador original de la Ingreso en el sistema transaccional
linea_detalle_proveedor	int	No nulo	Número de línea en la factura
fecha_id	int	 Llave foránea 	Llave foránea de DimFecha

articulo_id	int	Llave foránea	Llave foránea de DimArticulo
bodega_id	int	Llave foránea	Llave foránea de DimBodega
proveedor_id	int	Llave foránea	Identificador de DimProveedor
fecha_hora_transaccion	datetime	No nulo	Fecha y hora de registro de transacción
cantidad	decimal(28,13)	No nulo	Cantidad de unidades adquiridas
precio_total	decimal(28,13)	No nulo	Costo total en moneda local para la línea de Ingreso
costo_unitario	decimal(28,13)	No nulo	Costo unitario calculado como precio_total / cantidad

TABLA: DimProveedor

Campo	Tipo	Validación	Descripción
proveedor_key	int	AutoincrementalLlave primaria	Llave surroga de la tabla dimClientes
proveedor_id	varchar(20)	No nulo	Llave primaria del sistema transaccional
NIT	varchar(20)	No nulo	Numero de identificacion tributaria del cliente
nombre_proveedor	varchar(150)	No nulo	Nombre de la persona jurídica o natural
alias	varchar(150)	Default: N/A	Nombre con el que es conocida la empresa
fecha_ingreso	date	No nulo	Ubicación donde se encuentra el cliente
direccion	varchar(250)	Default: N/A	Nombre del contacto del cliente

contacto	varchar(30)	No nulo	Nombre del cargo del contacto del cliente
cargo_contacto	varchar(30)	No nulo	Correo electronico del cliente
pais	varchar(40)	No nulo	Primera opcion de telefono de contacto
condicion_pago	varchar(40)	No nulo	Segunda opcion de telefono de contacto
estado_proveedor	varchar(2)	No nulo	Nombre del país de origen del cliente
moneda	varchar(4)	No nulo	Nombre de la categoría del cliente asignada por la empresa.
telefono	varchar(20)	Default: No posee numero de contacto	Fecha en que entra en vigencia el registro
local	varchar(10)	Acepta nulos	Fecha en que finaliza la vigencia del registro
email	varchar(256)	Default: No posee email	
tipo_proveedor	varchar(40)	Default: SI	Variable que indica si es un registro vigente
fecha_inicio_vigencia	date	Default: fecha actual	Fecha en que entra en vigencia el registro
fecha_fin_vigencia	date	Acepta nulos	Fecha en que finaliza la vigencia del registro
registro_vigente	varchar(2)	Default: SI	Variable que indica si es un registro vigente

TABLA: DimBodega

Campo	Tipo	Validación	Descripción
bodega_key	int	AutoincrementalLlave primaria	Llave surroga de la tabla dimBodega
bodega_id	varchar(50)	No nulo	Llave primaria del sistema transaccional
bodega_nombre	varchar(50)	 No nulo 	nombre de la bodega
bodega_tipo	varchar(50)	No nulo	tipo de bodega
bidega_telefono	varchar(50)	Default: N/A	contacto de la bodega
bodega_direccion	text	Default: N/A	dirección de la bodega
bodega_traslados	text	Default: N/A	ubicación de la bodega
fecha_de_registro	date	 No nulo 	registros de la bodega
fecha_de_creacion	date	No nulo	registro de ingresos creado
produc_exitencia	varchar(120)	Default: N/A	existencia de los productos
zona_bodega	varchar(120)	No nulo	ubicación del traslado
codigo_cliente_consigna	varchar(120)	No nulo	codigo del cliente en consignacion
tipo_de _cambio	varchar(120)	Default: N/A	devoluciones
local	varchar(120)	No nulo	nombre del lugar

localización	varchar(120)	 No nulo 	ubicación y localización de producto	
--------------	--------------	-----------------------------	--------------------------------------	--

TABLA: FactTraspaso

Campo	Tipo	Validación	Descripción
traspaso_key	int	AutoincrementalLlave primaria	Llave subrogada de la tabla
fecha_id	int	Llave foranea	Llave foránea de DimFechas
articulo_id	int	Llave foránea	Llave foránea de DimArticulo
lote_id	int	Llave foránea	Llave foránea de DimLote
bodega_id	int	Llave foránea	Llave foránea de DimBodega
traspaso_id	int	Llave primaria	Llave primaria del sistema transaccional
fecha_hora_transaccion	datetime	 No nulo 	Momento en que se realiza la transacción
tipo	varchar(30)	No nulo	tipo de transacción venta o Ingreso
naturlaeza	varchar(30)	No nulo	naturaleza entrada o salida
fecha_registro	date	No null	fechas de los registros ingresados
cantidad	decimal(28,13)	No nulo	cantidad ingresada de producto

existencia_minima	decimal(28,13)	No nuloDefault: 0	Existencia mínima del producto
existencia_maxima	decimal(28,13)	No nuloDefault: 0	Existencia máxima del producto
cantidad_disponible_bruta	decimal(28,13)	No nulo	Cantidad disponible antes del consumo
cantidad_disponible_neta	decimal(28,13)	No nulo	Cantidad disponible después del consumo
costo_unitario	decimal(28,13)	No nulo	costo del articulo a traspasar
costo_total	decimal(28,13)	No nulo	costo total de los artículos

TABLA: DimArticulo

Campo	Tipo	Validación	Descripción
Articulo_key	int	AutoincrementalLlave primaria	Llave primaria subrogada
articulo_id	varchar(20)	No nulo	describe la llave primaria de dimensión articulo
descripcion	varchar(254)	No nulo	Es el nombre del articulo
tipo	varchar(20)	No nulo	Es el tipo producto de fabricación "Terminado" o "kit" esto por el proveedor
origen	varchar(20)	No nulo	Es el origen del artículo si de "Tercero" u otro

activo	varchar(20)	No nulo	Es el el estado del artículo "Activo" o "inactivo"
clasificacion1	varchar(40)	No nulo	Es la clasificación acorde al articulo esto se establece para identificar
clasificacion2	varchar(40)	No nulo	Es la clasificación acorde al articulo esto se establece para identificar
clasificacion3	varchar(40)	 No nulo 	Es la clasificación acorde al articulo esto se establece para identificar
clasificacion4	varchar(40)	No nulo	Es la clasificación acorde al articulo esto se establece para identificar
clasificacion5	varchar(40)	● No nulo	Es la clasificación acorde al artículo esto se establece para identificar
clasificacion6	varchar(40)	● No nulo	Es la clasificación acorde al articulo esto se establece para identificar
unidad_medida	varchar(6)	No nulo	Es la unidad de medida acorde al articulo
peso_neto	decimal(28,8)	No nulo	Describe el detalle del peso de articulo si aplica
peso_bruto	decimal(28,8)	No nulo	Describe el detalle del peso Bruto de articulo si aplica
volumen	decimal(28,8)	No nulo	Es la descripción del volumen del articulo si aplica
costo_local	decimal(28,8)	No nulo	Describe el costo del articulo es una operación de costo fiscal,coto STD y costo promedio

tipo_costo_comparativo	decimal(28,8)	No nulo	Es el tipo sistema de calificación de inventario "PEPS, promedio o estándar"
costo_comparativo	decimal(28,8)	No nulo	Es el ultimo costo de acorde al método de calculo
existencia_maxima	decimal(28,8)	No nulo	Es cantidad maxima que debe tener en stock el articulo
existencia_minima	decimal(28,8)	No nulo	Es cantidad mínima que debe tener en stock el articulo
punto_orden	decimal(28,8)	No nulo	Es la cantidad justa para que se orde una compara del articulo
ultima_salida	date	No nulo	es la fecha que describe la ultima salida que tuvo el articulo
ultima_ingreso	date	No nulo	es la fecha que describe el ingreso de un articulo al inventario
ultimo_inventario	date	 No nulo 	Es la fecha en la se realizó un control de inventario en caso de distribuidora es en bodega en caso de sala es por la encargada de los inventarios.

TABLA: DimVentas

Campo	Tipo	Validación	Descripción
ta lini	int	 Autoincremental 	Llave subrogada de la tabla
venta_key		Llave primaria	Liave Subiogada de la tabla
fecha_id	int	No nulo	Identificador de DimFecha
articulo_id	int	No nulo	Identificador de DimArticulo
bodega_id	int	No nulo	Identificador de DimBodega
cliente_id	int	No nulo	identificador de Dimcliente
lote_id	int	No nulo	Identificador de DimLote
venta_id	int	No nulo	Id del venta
detalle_venta_id	int	No nulo	Id del detalle de la venta
fecha_hora_transaccion	date	No nulo	Fecha hora de la transaccion
tipo_movimiento	varchar(15)	No null	Tipo de movimiento de la bodega
cantidad	decimal(13,8)	No nulo	Cantidad de producto a descontar
precio_local	decimal(13,8)	No nulo	Valor por la cual se vendio el producto precio es por unidad
costo	decimal(13,8)	No nulo	Costo del articulo

existencia_minima	decimal(13,8)	No nulo	Cantidad minima que debe tener la bodega
existencia_maxima	decimal(13,8)	No nulo	Cantidad maxima que debe tener la bodega
costo_articulos_vendido	decimal(13,8)	No nulo	Cantidad de un articulo antes del consumo
cantidad_disponible_neta	decimal(13,8)	No nulo	Valor calculado de cantidad_disponible_bruta - cantidad
costo_por_venta	decimal(13,8)	No nulo	es el costo de la venta de inventario unidades * costo
precio_total	decimal(13,8)	No nulo	es el precio de la venta unidades * precio local
margen_venta_dinero	decimal(13,8)	No nulo	es el margen de venta precio venta - costo venta
margen_porcentual	decimal(13,8)	No nulo	es el margen de venta (precio venta - costo venta)/precio venta