

Evidencia de aprendizaje 1: Modelo estrella de un Data Mart

Estudiantes:

**Elizabeth Alzate Murillo
Jimmy Antonio Mora Russy
Vanessa Foronda Moreno**

Institución: IU Digital de Antioquia

Ingeniería Software y Datos

Asignatura: Bases de Datos II

PREICA2401B010093

PREICA2401B010076

Docente: Víctor Hugo Mercado

21 abril 2024



Objetivo

Construir un modelo estrella para un *data mart* que permita analizar y responder a tres categorías específicas: identificar el producto más vendido, la categoría con más productos y el año con más ventas, utilizando la base de datos Jardinería como fuente de datos.

Instrucciones:

Sigue los pasos a continuación.

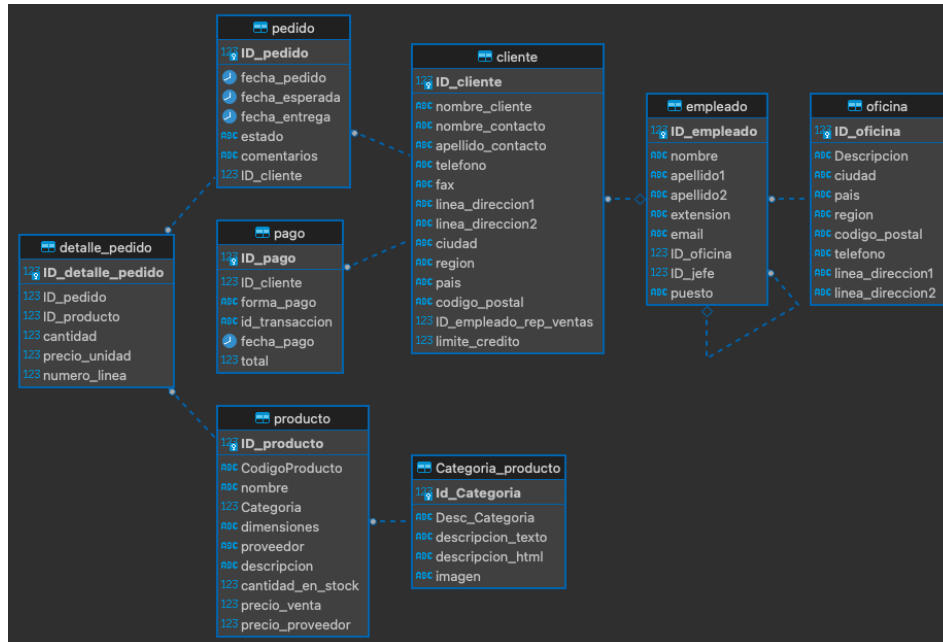
1. Análisis de la base de datos Jardinería, [consúltala aquí](#).
 - Revisa la estructura de la base de datos Jardinería para identificar las tablas relevantes y sus relaciones.
 - Identifica los campos necesarios para construir el modelo estrella, incluyendo la tabla de hechos y las dimensiones pertinentes.
 2. Diseño del modelo estrella:
 - Diseña la estructura tabla de hechos que represente las ventas o transacciones de la empresa.
 - Identifica y diseña las dimensiones relevantes que se relacionarán con la tabla de hechos.
 - Diseña las relaciones entre la tabla de hechos y las dimensiones, asegurando la integridad referencial.
 3. Documentación y presentación
- Redacta un documento que cumpla con normas APA y contenga:
- Introducción
 - Objetivos
 - Planteamiento del problema
 - Análisis del problema

- Propuesta de la solución con:
 - Descripción del modelo estrella propuesto.
 - Diseño (Imagen) del modelo estrella donde se puedan observar las dimensiones, la tabla de hechos, sus campos, tipos de datos y relaciones.
 - Lista de dimensiones propuestas.
 - Cada dimensión debe tener especificadas las columnas y qué tipos de datos van a almacenar.
 - Detalla la tabla de hechos, con sus campos y tipos de datos.
 - Conclusiones.
 - Anexos: puedes agregar el link donde construiste el modelo, también una imagen de mejor calidad del modelo.
 - Bibliografía.

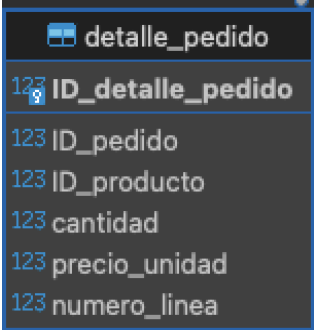
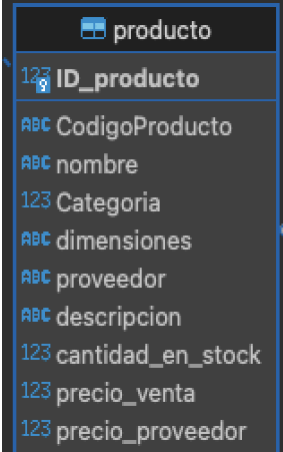
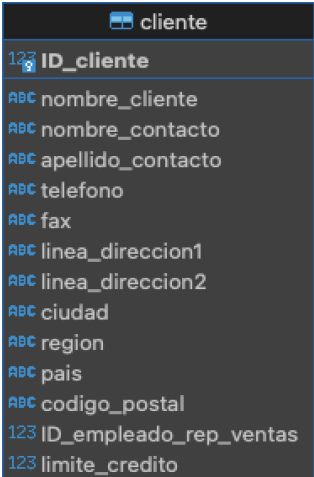
Desarrollo

1. Análisis de la base de datos Jardinería, [consúltala aquí](#).
 - a. Revisa la estructura de la base de datos Jardinería para identificar las tablas relevantes y sus relaciones.

Base de datos: Jardinería



Realizando la revisión de la estructura de la base de datos Jardinería identificamos las siguientes tablas relevantes y sus relaciones:

Tabla relevante	Tipo de tabla	Relaciones
	Transaccional / hechos	PK ID_Detalle_pedido FK ID_Producto FK ID_Pedido
	Dimensión	PK ID_Producto FK Id_Categoria_producto FK Id_detalle_pedido
	Dimensión	PK ID_cliente FK Id_Empleado FK Id_Pedido FK Id_Pago

<div><div>oficina</div><div><div>ID_oficina</div><div>Descripcion</div><div>ciudad</div><div>pais</div><div>region</div><div>codigo_postal</div><div>telefono</div><div>linea_direccion1</div><div>linea_direccion2</div></div></div>	Dimensión	PK ID_oficina FK Id_Empleado
---	-----------	-------------------------------------

- b. Identifica los campos necesarios para construir el modelo estrella, incluyendo la tabla de hechos y las dimensiones pertinentes.

Para construir el modelo estrella, consideramos las siguientes modificaciones al diagrama inicial:

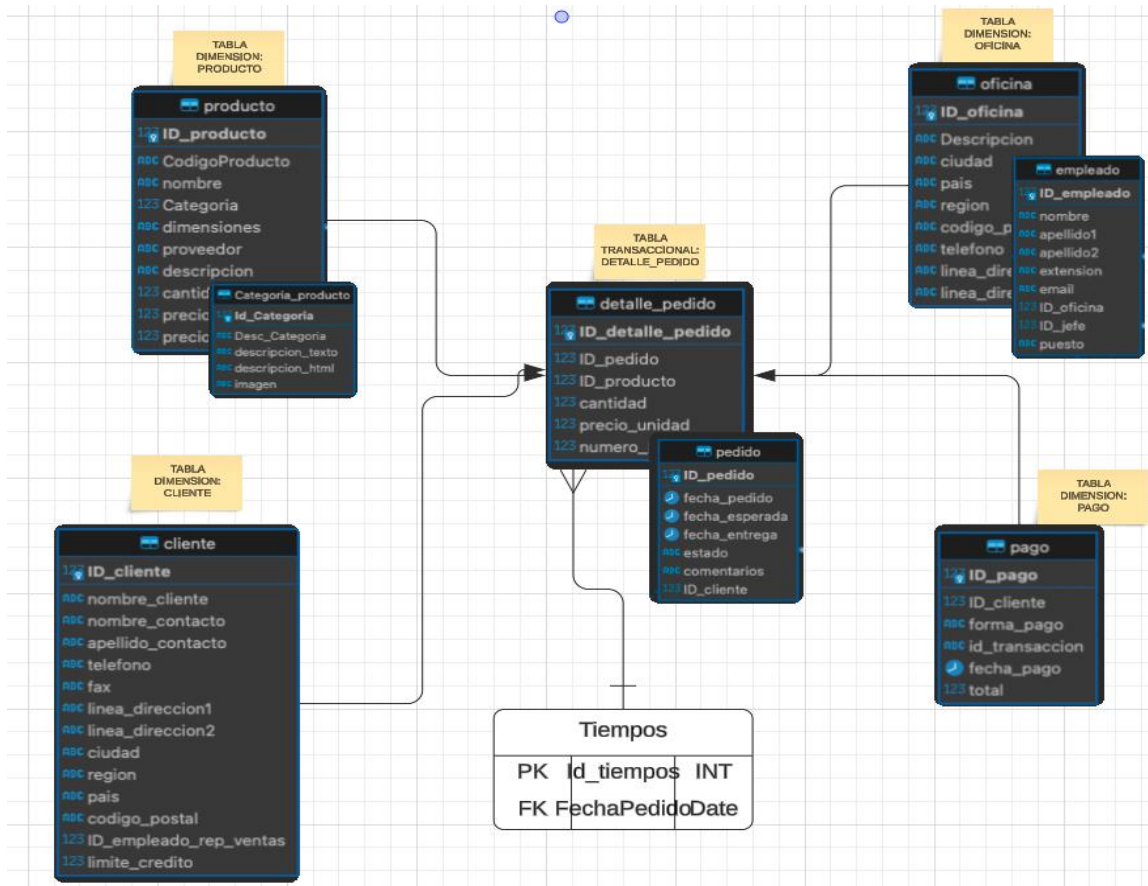
Nombre de la tabla	Tipo de tabla
Detalle de pedido - (Ventas)	Tabla Dimensional
Oficina (se une con tabla empleado)	Tabla Dimensional
Pago	Tabla Dimensional
Producto (Se une con tabla categoría de producto)	Tabla Dimensional
Cliente	Tabla Dimensional

Siendo así, se presentaron los siguientes cambios:

1. La tabla transaccional o de Hechos: "Detalle_Pedido" se unifica con la tabla pedido uniendo los ítems de las dos tablas. A futuro podrá ser nombrada tabla "ventas"
2. La tabla de dimensión Producto, acoge en unión a la tabla categoría productos, convirtiéndose en una sola tabla e incluyendo los campos de las dos tablas anteriores.
3. La tabla de dimensión oficina, se une con la tabla empleado, ya que consideramos que si bien es importante conocer y vincular una venta con un empleado, jerárquicamente la venta se hace a través de una oficina que tiene unos empleados. Siguiendo esta lógica consideramos importante para el modelo estrella unir ambas tablas.

4. Conservamos la tabla dimensional de pagos y clientes.

Diagrama de Dimensiones basados en nuestro modelo estrella:



2. Diseño del modelo estrella:

- a. Diseña la estructura tabla de hechos que represente las ventas o transacciones de la empresa.

detalle_pedido	
123	ID_detalle_pedido
	fecha_pedido
123	ID_cliente
123	ID_producto
123	cantidad
123	precio_unidad
123	numero_linea
	fecha_esperada
	fecha_Entrega
ABC	estado
ABC	comentarios

- b. Identifica y diseña las dimensiones relevantes que se relacionarán con la tabla de hechos

Tabla Dimensión Producto: Se une con la tabla Categoría.

producto	
123	ID_producto
ABC	CodigoProducto
ABC	nombre
ABC	dimensiones
ABC	proveedor
ABC	descripcion
123	cantidad_en_stock
123	precio_venta
123	precio_proveedor
123	Id_Categoria
ABC	Descrip_Categoria
ABC	Imagen

Tabla Dimensión Pago continúa igual, solo cambia su relación a la tabla de clientes a la tabla de detalle pedido.

pago	
123	ID_pago
ABC	forma_pago
ABC	id_transaccion
	fecha_pago
123	total
123	ID_detalle_pedido

Tabla dimensión oficina: se une con la tabla dimensión empleado, su relación se establece entre tabla de hechos detalle pedido y oficina.

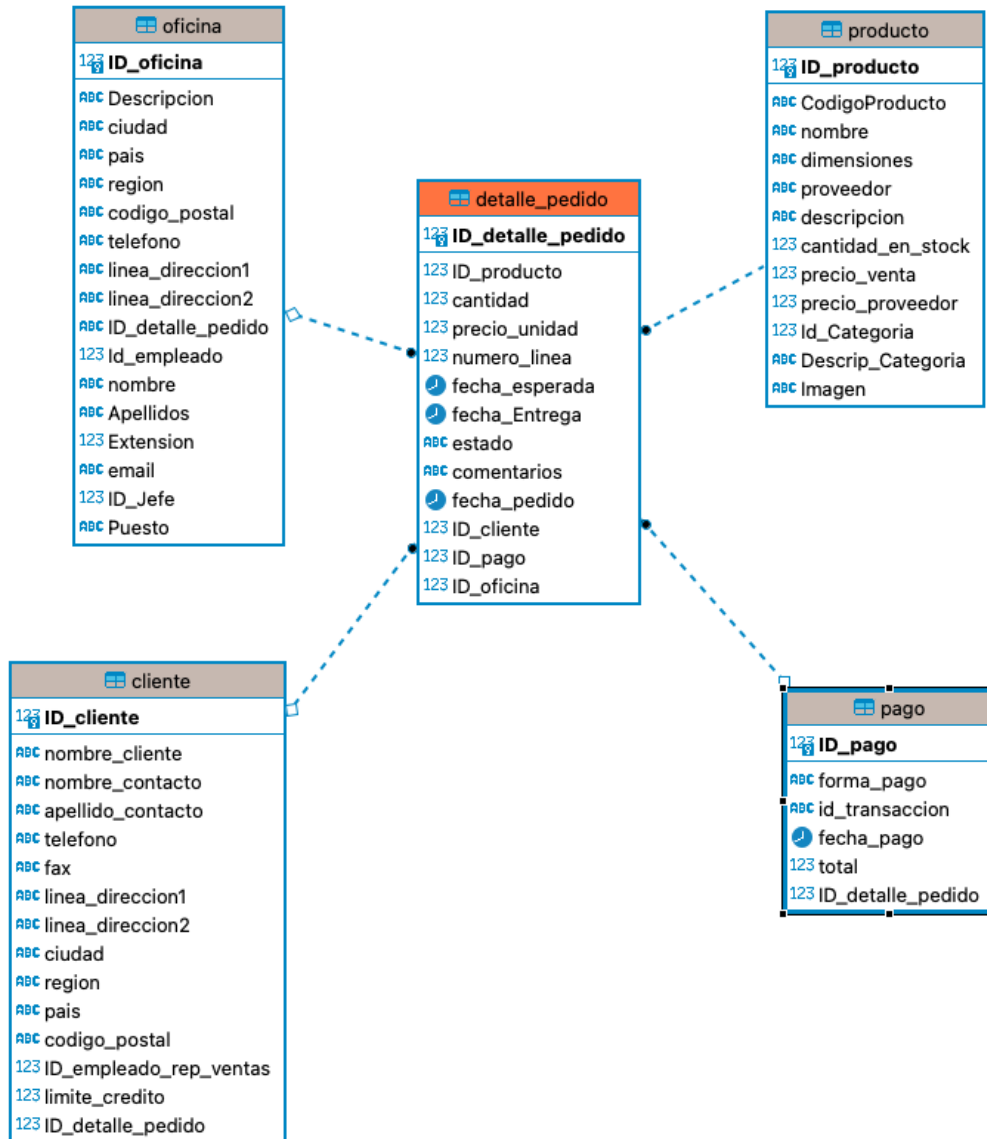
oficina	
123	ID_oficina
ABC	Descripcion
ABC	ciudad
ABC	pais
ABC	region
ABC	codigo_postal
ABC	telefono
ABC	linea_direccion1
ABC	linea_direccion2
ABC	ID_detalle_pedido
123	Id_empleado
ABC	nombre
ABC	Apellidos
123	Extension
ABC	email
123	ID_Jefe
ABC	Puesto

Tabla de dimensión cliente: Se conserva información de la tabla, se cambia la relación entre tabla pedido a la tabla de detalle pedido.

cliente	
123	ID_cliente
ABC	nombre_cliente
ABC	nombre_contacto
ABC	apellido_contacto
ABC	telefono
ABC	fax
ABC	linea_direccion1
ABC	linea_direccion2
ABC	ciudad
ABC	region
ABC	pais
ABC	codigo_postal
123	ID_empleado_rep_ventas
123	limite_credito
123	ID_detalle_pedido

c. Diseña las relaciones entre la tabla de hechos y las dimensiones, asegurando la integridad referencial.

A continuación, el diagrama de Relaciones entre la tabla de hecho las dimensiones, Se asegura Integridad referencial:



3. Documentación y presentación

Redacta un documento que cumpla con normas APA y contenga:

a. Introducción

En el ámbito actual, la analítica de datos y la utilización la de la información en la inteligencia de negocios, se han convertido en un precedente para las empresas que basan su estrategia en los datos, en la historia que cuentan, cómo se operaron y qué resultados se obtuvieron; de manera que, el aprendizaje obtenido durante este tiempo les permita aprovechar esa información y convertirla en una ventaja competitiva. Es allí donde radica la relevancia de los data marts, ya que brindan eficiencia en la recuperación de los datos, descentralizando la información y controlándola, permitiendo su análisis y extracción, convirtiendo datos rústicos en información útil para la toma de decisiones. A continuación, encontrará un ejercicio académico donde se han analizado datos de una base de datos y se han estructura de manera tal que sea un data marts confiable y efectivo para la actividad de la compañía, basados en un modelo estrella que permite responder a tres categorías específicas: identificar el producto más vendido, la categoría con más productos y el año con más ventas

b. Objetivos

Objetivo General: Construir un modelo estrella para un data mart que permita analizar y responder a tres categorías específicas: identificar el producto más vendido, la categoría con más productos y el año con más ventas, utilizando la base de datos Jardinería como fuente de datos.

Objetivos Específicos

- Identificar las dimensiones y hechos de la base de datos
- Plantear un modelo intuitivo y simplificado para el modelamiento de los datos
- Proponer una solución eficiente en la distribución de la información
- Suministrar un modelo operativo para la base de datos y sus transacciones
- Optimizar la estructura de la base de datos
- Facilitar la consulta y extracción de información para un análisis descriptivo de diagnóstico, predictivo o prescriptivo

c. Planteamiento del problema

Para la actividad, se emplea la base de datos llamada jardinería, en la que nos provee un modelo relacional basado en el cliente, en sus pedidos, tipo de pago y punto de venta inscrito. Asimismo, esta base contiene datos que no están relacionados en otras tablas (Ejemplo: proveedor), y otros como el dato imagen, que transaccionalmente no representa una referencia valiosa para su manipulación; por tanto, no provee un detalle sobre su relación o que suministre información importante en la dimensión asociada.

Siendo así, la compañía tiene el problema que sus datos no se encuentran estructuralmente organizados en un data mart, de tal forma que puedan extraer fácil y confiablemente información relevante a sus objetivos de identificar el producto más vendido, la categoría con más productos y el año con más ventas.

d. Análisis del problema

La base de datos jardinería propone el modelo de una base de datos relacional, en la que se destacan el hecho detalle pedido y la dimensión cliente, y a su vez otras dimensiones como pedido, oficina, empleado y producto, que se relacionan con las mencionadas anteriormente, brindando diferentes niveles de jerarquía, lo que hace que su acceso y manipulación se limiten a la jerarquía de la estructura, siendo un poco más complejo y con un mayor nivel de granularidad, posibilitando la inconsistencia y redundancia de los datos. Por lo tanto, se halla la oportunidad en mejorar la integridad referencial entre las tablas, asegurando la calidad de los datos y la relación entre las tablas con la creación del data mart en modelo estrella, que facilita el crecimiento incremental en el almacenamiento de los datos, debido a que se pueden añadir más tablas de hechos o dimensiones según la necesidad que se presente.

e. Propuesta de la solución con:**i. Descripción del modelo estrella propuesto.**

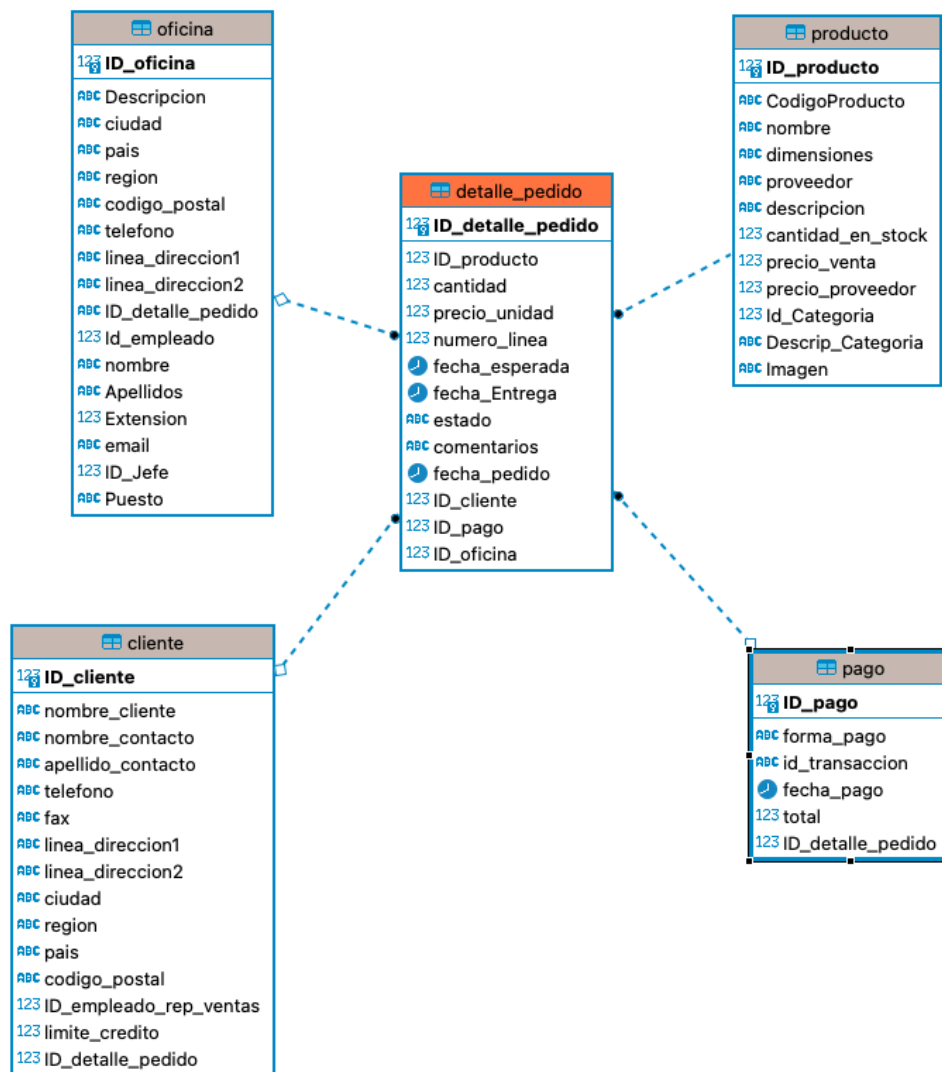
Nuestro modelo estrella propuesto reúne en una tabla central todos los hechos / transacciones que se pueda dar en la actividad comercial de la compañía, en esta tabla se responden las preguntas del quién compró, qué compró, cómo pagó, y dónde lo hizo. Adicional permite visualizar detalles como las fechas de estos registros (hechos), cantidades compradas, precios entre otros detalles. Así mismo la tabla “detalles de pedido” que es nuestro centro permitirá responder directamente a preguntas como cuál es el producto más vendido, o el año con más ventas, e indirectamente podría indicar cual es la categoría de productos con más productos.

Nuestra tabla de hechos se conecta con cuatro tablas de dimensiones a través de las FK, las cuales responden preguntas a gran detalle de campos muy particulares, por ejemplo: la tabla oficina que responde directamente donde se compró desde la tabla de hechos, podría generar información sobre cuál fue el empleado de este pedido y todos sus datos de contacto.

De igual manera la tabla de dimensión producto podría detallar información respecto a los orígenes del producto como tal, categoría, proveedores, precio de compra son algunos de los ejemplos de la información que nos arroja esta tabla.

Consideramos que con este modelo estrella, se cumple con los requisitos de la compañía en la creación del data mart ya que se puede encontrar toda la información de manera organizada, categorizada y con la trazabilidad transaccional requerida.

- ii. Diseño (Imagen) del modelo estrella donde se puedan observar las dimensiones, la tabla de hechos, sus campos, tipos de datos y relaciones.



iii. Lista de dimensiones propuestas.

Tabla de dimensión: Cliente

Tabla de dimensión: Pago

Tabla de dimensión: Oficina

Tabla de dimensión: Producto

- iv. Cada dimensión debe tener especificadas las columnas y qué tipos de datos van a almacenar.

cliente	oficina	producto
123 ID_cliente	123 ID_oficina	123 ID_producto
ABC nombre_cliente	ABC Descripcion	ABC CodigoProducto
ABC nombre_contacto	ABC ciudad	ABC nombre
ABC apellido_contacto	ABC pais	ABC dimensiones
ABC telefono	ABC region	ABC proveedor
ABC fax	ABC codigo_postal	ABC descripcion
ABC linea_direccion1	ABC telefono	123 cantidad_en_stock
ABC linea_direccion2	ABC linea_direccion1	123 precio_venta
ABC ciudad	ABC linea_direccion2	123 precio_proveedor
ABC region	ABC ID_detalle_pedido	123 Id_Categoria
ABC pais	123 Id_empleado	ABC Descrip_Categoria
ABC codigo_postal	ABC nombre	ABC Imagen
123 ID_empleado_rep_ventas	123 Apellidos	
123 limite_credito	123 Extension	
123 ID_detalle_pedido	ABC email	
	123 ID_Jefe	
	ABC Puesto	

pago
123 ID_pago
ABC forma_pago
ABC id_transaccion
123 fecha_pago
123 total
123 ID_detalle_pedido

- v. Detalla la tabla de hechos, con sus campos y tipos de datos.

detalle_pedido
123 ID_detalle_pedido
123 ID_producto
123 cantidad
123 precio_unidad
123 numero_linea
123 fecha_esperada
123 fecha_Entrega
ABC estado
ABC comentarios
123 fecha_pedido
123 ID_cliente
123 ID_pago
123 ID_oficina

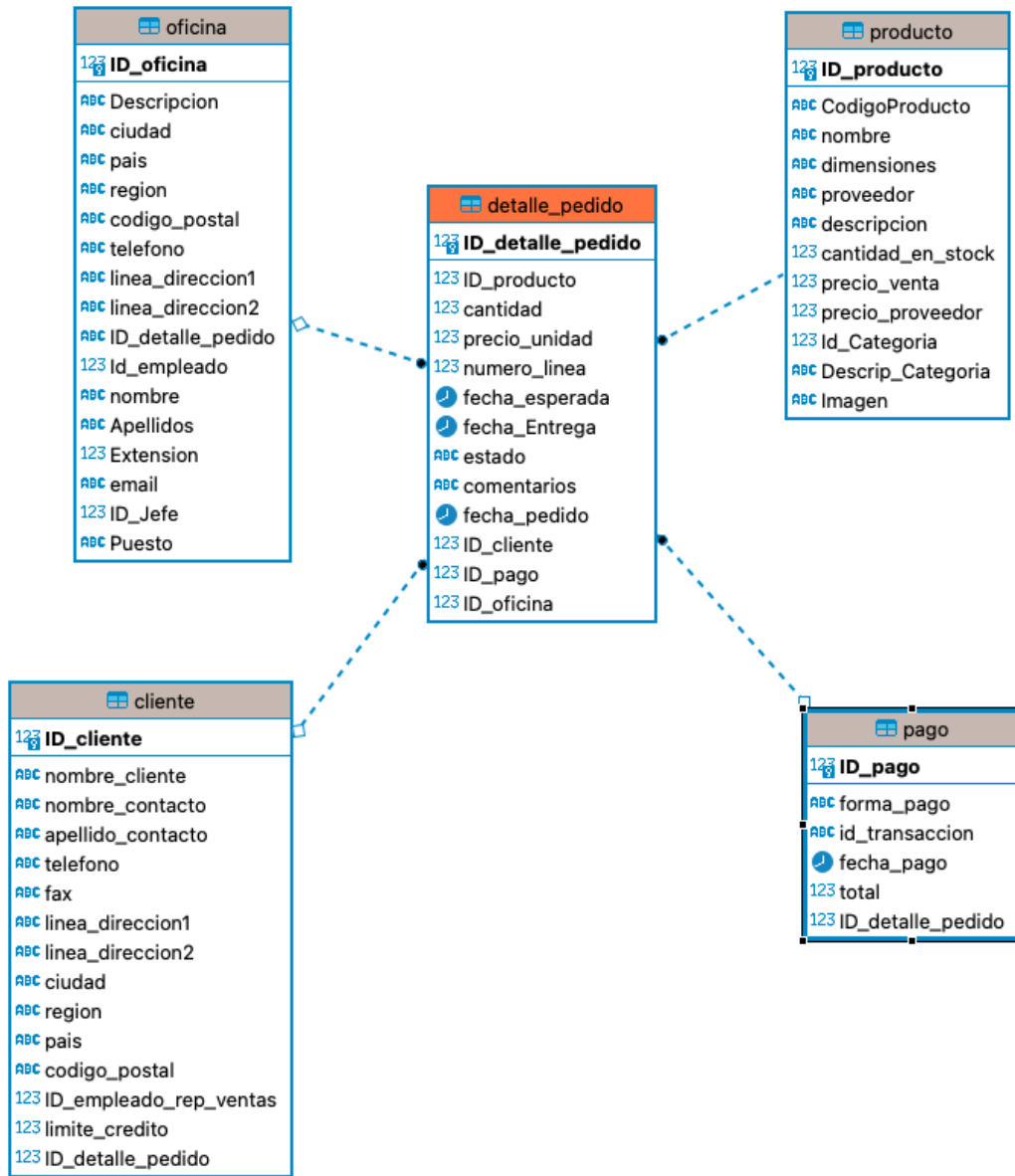
vi. **Conclusiones.**

Implementar un modelo estrella en los data marts para la base de datos jardinería, ha contribuido en una mejora significativa en la eficiencia de aquellas operaciones relacionadas a la misma, como las analíticas y transaccionales, facilitando la comprensión del negocio y los requisitos de los usuarios, colocando el lenguaje de estos datos en un lenguaje más familiar.

Al plantear una estructura simplificada, con una tabla central de hechos y las demás tablas como dimensiones (modelo Estrella), facilita a los usuarios finales la comprensión y el acceso a los datos necesarios para el análisis y toma de decisiones, adicional a que el modelo Estrella permitirá el escalamiento del almacenamiento de datos y la organización de los mismos de manera ordenada e incremental.

Asimismo, se podría integrar diferentes fuentes de datos con distintos niveles de detalle o granularidad, se destaca un notable rendimiento en las consultas tanto sencillas como complejas, permitiendo un mejor acceso a la información y trazabilidad de esta dentro del modelo.

- vii. Anexos: puedes agregar el link donde construiste el modelo, también una imagen de mejor calidad del modelo.



Bibliografía.

ComputerHoy.com [@Computerhoycom]. (2017, November 4). *¿Qué es Data Mining?* Youtube. <https://www.youtube.com/watch?v=ueAaIEroPY4>

Universidad de Deusto / Deustuko Unibertsitatea [@deusto]. (2015, November 26). *Introducción al Business Intelligence. Alex Rayón.* Youtube. https://www.youtube.com/watch?v=ElgUy_7eYLQ

Up, G. [@GrowUpDataAnalytics]. (2021, January 22). *¿Qué es Inteligencia de Negocios?* Youtube. <https://www.youtube.com/watch?v=ARr7RzdOLUI>