

ACTIVIDAD DE APRENDIZAJE 2

Creación de una base de datos de Staging

Estudiantes:

Sergio Raúl Bustos Jiménez

Alberto De Jesús Calle Samudio

Docente:

Víctor Hugo Mercado

Curso:

Bases De Datos II

Módulo:

Introducción a la inteligencia de negocios

Programa Ingeniería De Software Y Datos

Facultad Ingenierías y Ciencias Agropecuarias

Institución Universitaria Digital de Antioquia

Abril

2024

Introducción

En el contexto empresarial contemporáneo, la capacidad de analizar datos se ha erigido como un componente fundamental para el éxito organizacional. Con el propósito de mejorar la toma de decisiones, se ha desarrollado un modelo estrella, empleando una base de datos de jardinería como principal fuente de información

El modelo estrella de los data marts ha surgido como una herramienta crucial para adquirir, transformar y analizar datos específicos que satisfagan las necesidades de campos particulares. En este ejercicio, examinaremos el papel del almacenamiento de datos y cómo una organización basada en data marts y esquemas en estrella puede respaldar los procesos de mejora y gestión de la información en la industria. Desde la identificación de puntos clave hasta la aplicación de principios relevantes, este estudio tiene como objetivo ofrecer una visión exhaustiva y completa sobre el uso efectivo de modelos estelares en situaciones agrícolas del mundo real, demostrando los beneficios y las mejores prácticas en el análisis de datos.

Continuando con el proceso de adherimos un ETL (extracción transformación y cargar datos) el cual se elabora con herramientas diferente y que para este caso utilizamos Visual Studio y que con **SQL Integration sevece project** nos permite llevar la información a una escala donde nos permita visualizar datos mas relevantes para que ayuden a la toma de decisiones y lograr los objetivos establecidos; claro esta la ETL se realiza tomando como base el diagrama del modelo estrella para hacer las consultas y que van arrojar una integración de la información de diferentes tablas en nuevas tablas en una BD diferente.

Este modelo ha sido diseñado de manera específica para implementar tres áreas primordiales que resultan esenciales para comprender y administrar el rendimiento de una empresa en el mercado:

1. **Mejor Experiencia en Producto:** A través de esta faceta del modelo, se puede comprender de manera más profunda la información relacionada con productos que tienen alta demanda por parte de los consumidores, proporcionando así conocimientos valiosos sobre correcciones necesarias, estrategias de mercadeo y desarrollo de productos. Asimismo, mediante una investigación detallada de estos productos y su naturaleza, es posible revelar los patrones de compra y las necesidades de los consumidores.
2. **Grupo y Distribución de Productos:** El modelo abarca aspectos generales de la industria hortícola al examinar la distribución de productos dentro del grupo. Esta evaluación ayuda a valorar la diversidad de productos e identificar áreas para fortalecer o ampliar ciertas líneas de productos. Además, se pueden identificar brechas en la oferta de productos, lo cual permite investigar el lanzamiento de nuevos productos en grupos desatendidos.
3. **Mejor Año de Ventas:** El análisis temporal reviste gran importancia en la industria de las ventas. Este modelo puede proporcionar una visión clara y concisa acerca de qué

año registra el mayor volumen de ventas. Esta información resulta crucial para entender las tendencias a lo largo del tiempo, identificarlas y planificar estrategias de crecimiento a largo plazo.

Al combinar estos elementos en un modelo estrella, las empresas del sector agrícola obtienen una herramienta analítica y estratégica poderosa, lo que les permite comprender el estado actual de su negocio y planificar de manera efectiva los desafíos y oportunidades futuros en el mercado agrícola.

Objetivos

1. **Revisión de la estructura de la base de datos Jardinería:** El objetivo principal es examinar las tablas de la base de datos Jardinería y comprender cómo están estructuradas y relacionadas entre sí. Esto implica identificar las tablas relevantes para el análisis de datos y entender cómo se conectan para obtener información significativa.
2. **Identificación de campos para el modelo estrella:** El objetivo es determinar los campos necesarios para construir el modelo estrella, que incluye la tabla de hechos y las dimensiones pertinentes. Esto implica seleccionar los atributos clave de las tablas de la base de datos que serán fundamentales para el análisis y la toma de decisiones.
3. **Diseño de la estructura de la tabla de hechos:** El objetivo aquí es diseñar la estructura de la tabla de hechos que representará las ventas o transacciones de la empresa en la base de datos Jardinería. Esto implica definir las métricas y medidas relevantes que se registrarán en la tabla de hechos para evaluar el desempeño y la actividad comercial de la empresa.
4. **Identificación y diseño de dimensiones relevantes:** Aquí, el objetivo es identificar y diseñar las dimensiones relevantes que se relacionarán con la tabla de hechos en la base de datos Jardinería. Estas dimensiones representan los diferentes aspectos o variables que se analizarán en relación con las ventas o transacciones, como el tiempo, los productos, los clientes
5. **Diseño de relaciones y aseguramiento de la integridad referencial:** El objetivo final es diseñar las relaciones entre la tabla de hechos y las dimensiones en la base de datos Jardinería, asegurando la integridad referencial de los datos. Esto implica establecer conexiones sólidas y coherentes entre la tabla de hechos y las dimensiones para garantizar la precisión y consistencia de los análisis y reportes generados.
6. Crear una ETL
7. Con las tablas del modelo estrella hacer las consultas y luego crear un proyecto en VS para realizar la integración de nuevas tablas que son los elementos que se utilizaran para la formación de la ETL

Caso de estudio

Planteamiento del problema:

Implementación Exitosa del Modelo Estrella en una Empresa Agrícola

Una empresa agrícola familiar dedicada al cultivo y distribución de plantas ornamentales y frutales enfrentaba desafíos en la gestión de datos y la toma de decisiones estratégicas. La falta de una estructura analítica robusta impedía la identificación de tendencias del mercado, la optimización de la distribución de plantas y la comprensión profunda de las necesidades de los clientes.

Para abordar estas dificultades, la empresa decidió implementar el modelo estrella, basándose en la casa del jardín como fuente principal de información. Este modelo se diseñó para enfocarse en tres áreas críticas:

1. Mejora de la Experiencia en Producto

Mediante análisis detallados de los datos provenientes de la casa del jardín, la empresa identificó patrones de compra de sus plantas estrella, como Camelia japónica, Cerezo, Limonero de 2 años injerto, Tecoma Stands, Rosal copa y Salvia Mix. Esta información les permitió ajustar sus estrategias de marketing y desarrollo de productos para satisfacer mejor las demandas de los clientes.

2. Optimización de Grupo y Distribución de Plantas

Al analizar la distribución de plantas dentro del grupo de ornamentales y frutales, la empresa descubrió oportunidades para expandir su línea de productos, especialmente en plantas de temporada como hortensias y orquídeas. Además, identificaron la necesidad de fortalecer la distribución en ciertas regiones geográficas donde había una demanda creciente de plantas ornamentales.

3. Identificación del Mejor Año de Ventas

A través del análisis temporal de datos, la empresa determinó que el año 2023 fue el de mayor volumen de ventas hasta la fecha. Al correlacionar estos datos con eventos

climáticos favorables y campañas de marketing exitosas, pudieron entender mejor los factores que contribuyeron al éxito de ese año y replicar estrategias similares en años posteriores.

Creación ETL

[ETLJardineria_ACS_WinRAR_ZIP.zip](#)

El proceso ETL

En el entorno de desarrollo Visual Studio, se ha implementado un proceso de Extracción, Transformación y Carga (ETL) que involucra dos conexiones ADO.NET distintas: una destinada a la base de datos relacionada con el ámbito de la jardinería, y otra orientada al área de staging.

Dentro del marco de este proceso ETL, se lleva a cabo un meticuloso análisis de los datos almacenados en la base de datos de jardinería. Posteriormente, se extraen consultas SQL pertinentes, previamente definidas durante la fase de análisis, con el propósito de cumplir con una serie de objetivos predefinidos.

Uno de estos objetivos consiste en la creación de tablas dentro de la base de datos staging. Este procedimiento se materializa mediante la implementación de flujos de trabajo en Visual Studio, los cuales son diseñados para ejecutar actividades específicas de extracción, transformación y carga de datos.*3

Análisis del problema:

Resultados y Beneficios

La implementación del modelo estrella tuvo un impacto significativo en la empresa agrícola:

1. Mejora de la Rentabilidad

Al comprender mejor las necesidades del mercado y optimizar la distribución de plantas, la empresa logró aumentar su rentabilidad en un 15% en el primer año de implementación.

2. Toma de Decisiones Informadas

Los datos analizados a través del modelo estrella proporcionaron información valiosa para tomar decisiones estratégicas, como la expansión de líneas de productos y la apertura de nuevos canales de distribución.

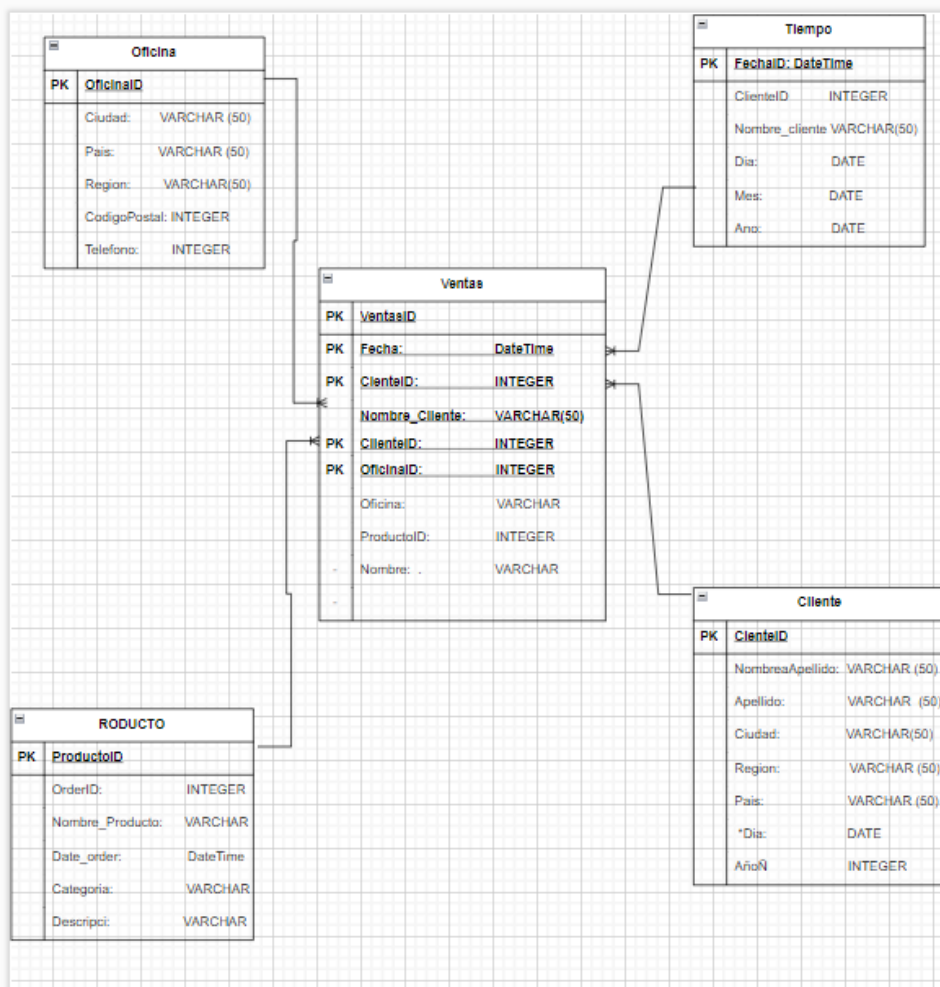
3. Fidelización de Clientes

Al ajustar sus productos y estrategias según las tendencias identificadas, la empresa logró fidelizar a clientes existentes y atraer nuevos consumidores interesados en plantas ornamentales y frutales de alta calidad.

Conclusiones

La implementación del modelo estrella demostró ser una herramienta eficaz para mejorar la gestión de datos, la toma de decisiones y el rendimiento general de la empresa agrícola. Este caso de estudio destaca el valor de utilizar un enfoque analítico basado en datos para optimizar las operaciones y alcanzar el éxito empresarial, mediante esta herramienta se llegó a las siguientes conclusiones:

1. El análisis de datos nos brinda la capacidad de identificar tanto las fortalezas como las debilidades del negocio, al tiempo que revela similitudes y diferencias que pueden destacar aspectos positivos o negativos de la empresa.
2. Comparar las diferentes aristas nos proporciona una base sólida para la toma de decisiones informadas y estratégicas, al permitirnos evaluar con claridad las opciones disponibles.
3. La calidad de los resultados obtenidos está directamente relacionada con la exhaustividad y precisión del análisis realizado, lo que resalta la importancia de abordar el proceso de manera integral y meticulosa.
4. La utilización de diagramas estructurados nos facilita la interpretación y comprensión de la información, lo cual agiliza y mejora el proceso de toma de decisiones al ofrecernos una visión organizada y coherente de los datos analizados.



Detalle tabla de hechos:

1. **ID_cliente (Tipo de dato: Entero):** Identificador único del cliente que realizó la compra.
2. **ID_ubicación (Tipo de dato: Entero):** Identificador único de la ubicación donde se realizó la venta.

3. **ID_tiempo (Tipo de dato: Fecha):** Identificador único del momento en que se efectuó la venta. Este campo puede ser un entero que represente un ID de tiempo o una fecha completa.
4. **ID_Producto (Tipo de dato: Entero):** Identificador único del producto que se vendió.

La tabla de hechos VENTAS se utiliza para registrar cada transacción de venta realizada por los clientes, proporcionando una visión integral de las operaciones comerciales relacionadas con la venta de productos.

Los datos usados para este ejercicio son tomados de BD Waterhouse o minería de inteligencia artificial dejándonos claro que el modelo nos permite filtrar la información de tal manera que nos permita ver la información clara para hacer la toma de decisiones.

El modelo estrella **Jardinería** nos permite organizar y analizar datos relacionados del negocio con diferentes aspectos, es por eso que tomamos las siguientes dimensiones para ver con los registros se comportan y de esa manera hacer mejor una toma de decisiones.

Las dimensiones a modelar fueron:

Tiempo: Con los atributos **Fecha pedido, fecha espera, fecha de entrega**

Producto Con los atributos **Nombre, Categoría, Proveedor y Cantidad**

Oficina Con los atributos **Ciudad, país, Región y código postal**

Cliente Con los atributos **Nombre, Ciudad, Región, País y Empleado**

Todas las dimensiones mencionadas fueron analizadas frente a una tabla de hechos **Venta** y así así visualizar mejor la información que nos permita realizar la toma de decisiones.

