

PROYECTO FINAL SQL
COMISIÓN 397 80
ALEJANDRO GIORGIO

1.- INTRODUCCIÓN

En el presente trabajo se plantea una base de datos de una distribuidora de cigarrillos denominada “Tobacco N’ Stuff”. En esta base de datos relacional se toma como la tabla “core” o “de hechos” la correspondiente a las ventas de la distribuidora. En esta tabla constan las distintas ventas mayoristas y minoristas así como tablas que dimensionen estas ventas por: clientes y sus dimensiones, tipo de descuentos, productos que componen la venta, distintos tipos de impuestos que se aplican según el impuesto en cuestión, etc.

El Diagrama Entidad Relación se plantea en formato “Copo de Nieve” al haber tablas de dimensiones que cuentan con una segmentación propia con sus propias dimensiones.

2.-OBJETIVO

El objetivo de la presente DB es generar un modelo que no solo optimice y traiga claridad a las cuentas de la empresa, tanto en la forma de almacenamiento de datos y el input de los mismos logrando una mayor eficiencia, sino también generar reportes de interés con los datos almacenados a fines de poder dar insights valiosos rápidamente generados para una mas ágil toma de decisiones.

3.- SITUACIÓN PROBLEMÁTICA

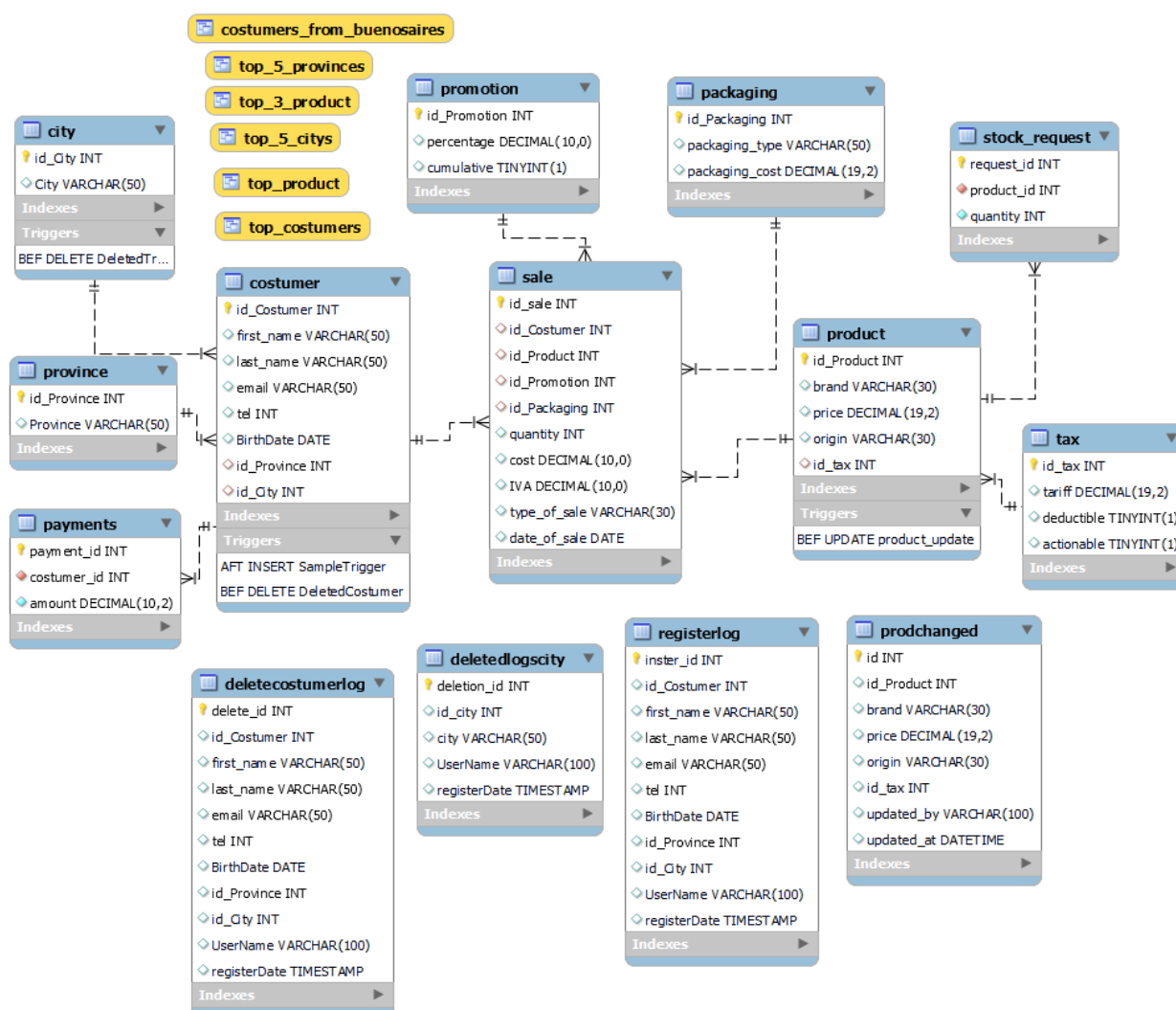
Se presenta la situación de que la distribuidora se desempeña en un mercado muy volátil, altamente informal y burocrático. Los impuestos aplicados sobre los productos del tabaco (muy distintos para cada producto), la naturaleza recurrible por precedentes jurisprudenciales sobre los mismos, el volumen de ventas, los ingresos y egresos de mercadería y la falta de uniformidad de todos estos factores entre diferentes marcas, genera una obscuridad sobre ciertos datos claves de la distribuidora la cual no cuenta con la infraestructura contable y de sistemas para afrontar este nivel de auditoria sobre sus transacciones diarias. Actualmente realizan escuetas anotaciones sobre documento papel y cuentan con un software donde se registran los ingresos brutos a fines de generar informes fiscales, mas no cuentan con certeza de cuanto es la ganancia neta, los mejores clientes, los mejores productos, las ciudades con mas ventas, etc etc.

4.- MODELO DE NEGOCIOS

El modelo de negocios se estructura al rededor de las ventas las cuales son el “core” de negocios. De alli se desprenden las demás dimensiones. La distribuidora hace venta mayorista de productos del tabaco a distintos clientes que hacen reventa minorista a lo largo de todo el país. Los mismos hacen transferencias a la cuenta recaudadora de la misma y retiran stock.

5.- DIAGRAMA DE ENTIDAD RELACIÓN

Se adjunta el DER del la distribuidora TobbacoNstuff. El mismo cuenta con 14 tablas. Las tablas que se encuentran en el margen inferior son tablas de auditoria o transaccionales las cuales no se encuentran directamente conectadas mediante cardinalidad al modelo entidad relación ya que tienen fines de auditoria y control.



2.- Listado de tablas

A continuación el listado de las tablas con su PK FK y su tipo de dato, junto con una breve descripción de cada tabla:

-Tabla de provincias con una PK única para cada provincia

TABLE Province

id_Province INT PRIMARY KEY

Province VARCHAR(50)

-Tabla de ciudades con una PK unica por ciudad

TABLE City

id_City INT PRIMARY KEY

City VARCHAR(50)

-Tabla de cliente. Por un error de tipeo que paso inadvertido la tabla quedo mal llamada “costumer” en lugar de “customer”. Contiene los datos descriptivos del cliente así como la FK que remite a la provincia y ciudad de origen

TABLE Costumer

id_Costumer INT PRIMARY KEY

first_name VARCHAR(50)

last_name VARCHAR(50)

email VARCHAR(50)

tel INT

BirthDate DATE

id_Province INT FK

id_City INT FK

-Tabla de promoción, donde se especifica que promoción esta disponible en una determinada venta y que porcentaje de descuento contiene esa promoción

TABLE Promotion

id_Promotion INT PRIMARY KEY

percentage DECIMAL
cumulative BOOL

-Tabla que describe el tipo de empaquetado y el costo del mismo correspondiente a cada venta

TABLE Packaging

id_Packaging INT PRIMARY KEY
packaging_type VARCHAR(50)
packaging_cost DECIMAL(19,2)

-Tabla que contiene información sobre el impuesto aplicado sobre la venta, si es deducible o si es apelable judicialmente

TABLE Tax

id_tax INT PRIMARY KEY
tariff DECIMAL(19,2)
deductible BOOL
actionable BOOL

-Tabla que contiene información sobre el producto en cuestión. Que marca es, que precio tiene, el origen del mismo y que impuesto se le aplica

Tabla de producto

TABLE Product

id_Product INT PRIMARY KEY
brand VARCHAR(30)
price DECIMAL(19,2)
origin VARCHAR(30)
id_tax INT FK

- La tabla de hechos. Contiene toda la información relevante para la venta en cuestión así también con la FK del cliente, producto, promoción y packaging

TABLE Sale

id_sale INT PRIMARY KEY,
id_Costumer INT FK

id_Product INT FK
id_Promotion INT FK
id_Packaging INT FK
quantity INT
cost DECIMAL
IVA DECIMAL
type_of_sale VARCHAR(30)
date_of_sale DATE

-Tabla de auditoria activada por un trigger que mantiene un control de los registros de ingreso de clientes con todos sus datos, así como usuario y timestamp del momento en que se realizo un insert

TABLE registerLog

inster_id INT PRIMARY KEY
id_Costumer INT
first_name VARCHAR(50)
last_name VARCHAR(50)
email VARCHAR(50)
tel INT
BirthDate DATE
id_Province INT
id_City INT
UserName varchar(100) NULL
registerDate timestamp

-Tabla de auditoria activada por trigger que deja constado si una ciudad fue borrada, por que personal y en que momento

TABLE DeletedLogscity

deletion_id INT NOT NULL AUTO_INCREMENT PRIMARY KEY
id_city int
city varchar(50)
UserName varchar(100) NULL
registerDate timestamp

-Contraparte de "registerlog". Tabla de auditoria que se activa por un trigger y mantiene un control de los registros borrados de clientes con

todos sus datos, así como usuario y timestamp del momento en que se realizó un delete

TABLE DeleteCostumerLog

delete_id INT PRIMARY KEY
id_Costumer INT
first_name VARCHAR(50)
last_name VARCHAR(50)
email VARCHAR(50)
tel INT
BirthDate DATE
id_Province INT
id_City INT
UserName varchar(100) NULL
registerDate timestamp

-Tabla de auditoria activada por trigger que registra los productos que fueron actualizados dejando registro de los como se encontraban antes del update.

TABLE ProdChanged

id INT PRIMARY KEY,
id_Product INT,
brand VARCHAR(30),
price DECIMAL(19,2),
origin VARCHAR(30),
id_tax INT,
updated_by VARCHAR(100),
updated_at DATETIME,

-Tabla transaccional en la cual se ingresan datos a través de un store procedure que asegura la integridad de la transacción. Esta tabla contiene los pagos de los clientes junto con su importe y una identificación de cada pago

TABLE payments

payment_id INT AUTO_INCREMENT PRIMARY KEY
costumer_id INT NOT NULL

amount DECIMAL(10,2) NOT NULL

-Tabla transaccional en la cual se ingresan datos a través de un store procedure que asegura la integridad de la transacción. Constan en la misma los pedidos de stock de determinados productos. Con una clave identificatoria para cada pedido.

TABLE stock_request

request_id INT KEY

product_id INT NOT NULL

quantity INT NOT NULL

3.-SCRIPT SQL

Dentro del repositorio se encuentran los siguientes scripts

LINK: https://github.com/AlejandroGiorgio/SQL_CH.git

- DDL con la creación de la base de datos: Create_tobbacoNstuff.sql
- Inserción de datos a la base de datos: Insert_tobbacoNstuff.sql
- Funciones de la DB: Functions_tobbacoNstuff.sql
- Store procedures de uso general: StoredProcedure_tobbacoNstuff.sql
- (OPCIONAL) Prueba de DCL: Tobbaconstuff_DCL.sql
- Creación de tablas transaccionales y TCL: Tobbaconstuff_TCL.sql
- Creación de 5 informes de la DB: View_tobbacoNstuff.sql
- Creación de triggers y tablas de auditoria: tobbacoNstuff_triggers.sql

5.-INFORMES GENERADOS:

top_5_Provinces: Esta vista muestra las 5 provincias con más pedidos de clientes. Selecciona el id de la provincia y cuenta el número de pedidos por cliente en esa provincia. Se utiliza para mantener una revision constante de las provincias con mas ventas

Costumers from BuenosAires: Esta vista muestra todos los clientes de la provincia con id 1 (Buenos Aires). Siendo la provincia con mas habitantes del pais sirve para mantener un monitoreo del numero de clientes de esa provincia

Top_Costumers: Esta vista muestra los 10 clientes con más pedidos. Selecciona el id del cliente, su nombre completo y cuenta el número de pedidos por cliente para cruzar estos datos con los demás informes y generar insights sobre esta temática.

Top_3_product: Esta vista muestra los 3 productos más vendidos. Selecciona el id del producto y cuenta el número de ventas por producto para conocer cual es el producto que mas se ha vendido.

top_5_citys: Esta vista muestra las 5 ciudades con más pedidos de clientes. Selecciona el id de la ciudad y cuenta el número de pedidos por cliente en esa ciudad.

6.-Herramientas y tecnologías usadas

Para este trabajo se utilizo principalmente MYSQL Workbench para la creación de la DB y la inserción de los datos.

Los datos fueron generados de forma ficticia utilizando Python con listas de datos ficticios y la librería “random”.

7.-Futuras

lineas

A partir de este modelo relacional se espera poder implementar a futuro algoritmos de regresión lineal y modelos de machine learning para detectar tendencias de consumo por parte de los clientes y poder realizar análisis predictivos y prescriptivos sobre decisiones comerciales provechosas para la TobbacoNstuff.