**PROYECTO FINAL**

**CARRERA DATA ANALYTICS** 

**Ignacio Brunotto**

**INTRODUCCIÓN**

**"Análisis de datos y tendencias de consumo en la base de datos de la supertienda: Un enfoque hacia la optimización de la estrategia de ventas y fidelización de clientes"**

**1. Descripción de la temática de los datos**

Para el siguiente trabajo, se extrajo información acerca de una tienda con ventas masivas de los Estados Unidos.

Se debe tener en cuenta que el dataset contiene información tanto de los clientes, del producto y especificaciones sobre la compra realizada (ship date, order date, etc).

Este dataset contiene datos que van desde el 2014 hasta el 2017

**2. Hipótesis**

En un mercado donde hay muchos competidores creciendo y con una suba en la demanda, lo que se busca es entender que es lo que mejor se ajusta a lo que es mejor para la tienda. Definir en qué productos, categorías, regiones y segmentos de consumidores deberían enfocarse y con ello establecer estrategias acordes con los objetivos generales de la misma.



**1. DESCRIPCIÓN DE LOS DATOS**

Hoy en día las grandes tiendas pueden llegar a cualquier parte del país, a gran velocidad, debido a ello generar grandes movimientos en los flujos de dinero, favoreciendo así a la economía del mismo. Debido a esto, en el siguiente trabajo, se realizará un análisis de las ventas de los años 2011-2014 de una ‘super store’. Los datos fueron obtenidos de Kaggle: https://www.kaggle.com/datasets/bravehart101/sample-supermarket-dataset

**2. HIPÓTESIS DE TRABAJO**

La hipótesis en la que se basa el trabajo es en que hay distintas variables que afectan a las ventas, como puede ser el perfil demográfico del cliente, el método de envío y las diferentes categorías y subcategorías, y cómo estas varían según lo mencionado anteriormente.

**2.1 Resultados esperados**

- Aumento en las ventas comparado a los años anteriores

- Crecimiento en las ganancias por ventas en comparación a años anteriores

- Información valiosa acerca de los segmentos demográficos que generan los montos promedios de ventas más altos

- Información valiosa sobre cuáles son los segmentos, categorías y subcategorías que

que generan los promedios de ventas más altos

**3. OBJETIVO GENERAL**

Obtener información y conocimientos accionables que permitan tomar decisiones estratégicas para mejorar el rendimiento y la eficiencia de la tienda en línea. Esto implica explorar y comprender los patrones, tendencias y relaciones en los datos relacionados con las ventas, los clientes y otros aspectos relevantes del negocio en línea.

**3.1 Objetivos específicos**

- Comprender los patrones de ventas a lo largo del tiempo, identificar los productos

más vendidos, las categorías de mayor demanda y los momentos con mayor actividad comercial. - Segmentar y comprender a los clientes según diferentes criterios demográficos y comportamentales - Determinar la relación entre las ventas y las ganancias obtenidas

**4. ALCANCE**

Con base en los datos obtenidos en la tabla se busca determinar cuales son los puntos en donde debemos fortalecer nuestro negocio y que otros son nuestras mayores fortalezas

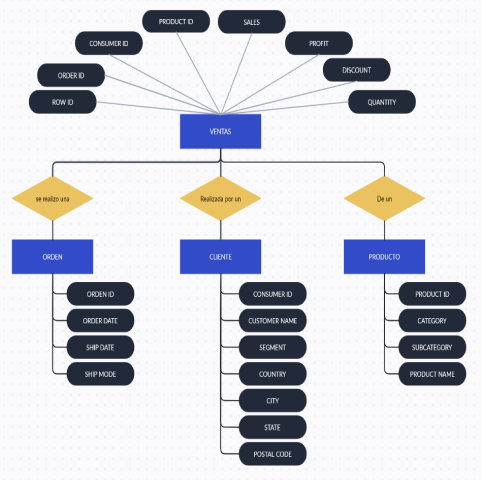
**5. USUARIO FINAL Y NIVEL DE APLICACIÓN DEL ANÁLISIS**

Los usuarios finales del dashboard serán los directivos y encargados de tomar decisiones sobre la conducción de la tienda. Por lo que el nivel de análisis será a niveles estratégicos en donde se buscará dar información útil y certera que facilite todo el proceso que implica la toma de una decisión

**6. DATASET**

El dataset fue obtenido de Kaggle https://www.kaggle.com/datasets/bravehart101/sample-supermarket-dataset

**7. DIAGRAMA ENTIDAD-RELACIÓN**

****

**8. LISTADO DE TABLAS**

A) TABLA VENTAS

B) TABLA ORDENES

C) TABLA DE CLIENTES

D) TABLA DE PRODUCTOS

**9. LISTADO COLUMNAS DE TABLAS**

**Tabla 1: TABLA DE VENTAS** En esta tabla encontraremos el código de la venta, junto al de producto, orden y cliente. Como así también los detalles específicos de cada compra (precio, cantidad, profit y descuento)

| **TIPO DE CLAVE** | **CAMPO TIPO DE CAMPO** | **DETALLE** |
| --- | --- | --- |
| PK | ROW ID INT | Código de identificación único de la venta |
| FK | ORDER ID VARCHAR | Código de identificación de la orden. Permite la  relación con la tabla  ORDEN |
| FK | PRODUCT ID VARCHAR | Código de identificación de los productos. Permite la relación la tabla  PRODUCTO |
| FK | CONSUMER ID VARCHAR | Código de identificación del cliente. Permite la relación con la tabla CLIENTE |
|  | SALES FLOAT | Monto por el cual fue  realizada la venta |
|  | QUANTITY INT | Cantidad de productos vendidos |
|  | PROFIT / FLOAT | Ganancia de la operación  realizada |
|  | DISCOUNT / FLOAT | Desscuentos aplicados a la  compra |

**TABLA 2 ‘ORDEN’:** En esta tabla encontraremos todo lo relacionado a la orden que fue realizada. Con esto nos referimos a la fecha en que se realizó el pedido, la fecha de envío y el modo de envío

| TIPO DE CLAVE | CAMPO / TIPO DE CAMPO | DETALLE |
| --- | --- | --- |
| PK | ORDER ID / VARCHAR  ORDER DATE / DATE | Codigo de identificacion de la orden, nos permite  relacionar esta tabla con la tabla VENTAS  Fecha en la que fue  realizada el pedido |
|  | SHIP DATE / DATE | Fecha de envío del pedido |
|  | SHIP MODE / VARCHAR | Modo en el que fue enviado el pedido |

**TABLA 3 ‘CLIENTE’:** En esta tabla encontraremos todo lo relacionado al cliente que realizo la orden. Datos demográficos, datos personales

| TIPO DE CLAVE | CAMPO / TIPO DE CAMPO | DETALLE |
| --- | --- | --- |
| PK | CONSUMER ID / VARCHAR | Código de identificación del cliente, nos permite relacionar esta tabla con la tabla VENTAS |
|  | CONSUMER NAME / VARCHAR | Nombre del consumidor que realizó la orde |
|  | SEGMENT /VARCHAR | Segmento del cual forma parte el  comprador, si es consumidor final o si compra para alguna empresa o realiza home office |
|  | COUNTRY / VARCHAR | País desde el cual se genera la orden, y a donde se realiza el envío |
|  | CITY / VARCHAR | Ciudad desde el cual se genera la orden, y a donde se realiza el envío |
|  | STATE / VARCHAR | Estado desde el cual se genera la orden, y a donde se realiza el envío |
|  | POSTAL CODE / INT | Código postal del lugar donde se hace el pedido |

**TABLA 4 ‘PRODUCTO’:** En esta tabla encontraremos todo lo relacionado al producto que fue comprado, categoría, subcategoría y su nombre.

| **TIPO DE CLAVE** | **CAMPO TIPO DE CAMPO** | **DETALLE** |
| --- | --- | --- |
| PK | PRODUCT ID VARCHAR  PRODUCT NAME VARCHAR | Codigo de identificacion de el producto, el cual permite relacionar esta tabla con la tabla VENTAS  Nombre del producto que fue comprado |
|  | CATEGORY VARCHAR | Categoría dentro de la cual se ubica el producto  comprado |

SUBCATEGORY VARCHAR Subcategoría dentro de la cual se encuentra ubicado el pedido que fue comprado

**Transformaciones realizadas**

-Cambio en tipo de datos en las columnas sales, profit y discount

-En la tabla productos, se modificó los títulos.

-Volver a eliminar duplicados en la tabla producto como en la tabla clientes, ya que no me dejaba realizar la relación.

-En la tabla ventas, se dividió la columna order ID para poder obtener el año de la venta

**Medidas creadas**

**1. Cantidad de operaciones:**

****

**2. Cantidad de productos:**

****

**3. Cantidad de estados:**

****

**4. Primera operación realizada:**

****

**5. Ventas totales**

****

**6. Ventas 2017**

****

**Columnas creadas**

**1. Ventas totales por operación:**

****

**2. Nivel de facturación:**

****