**Caso 1. Ventas de autos y motocicletas de colección**

Este conjunto de datos corresponde a las ventas de una tienda minorista. La información contiene pedidos, clientes, envíos, etc. Este conjunto hace parte originalmente e Pentaho DI Kettle.

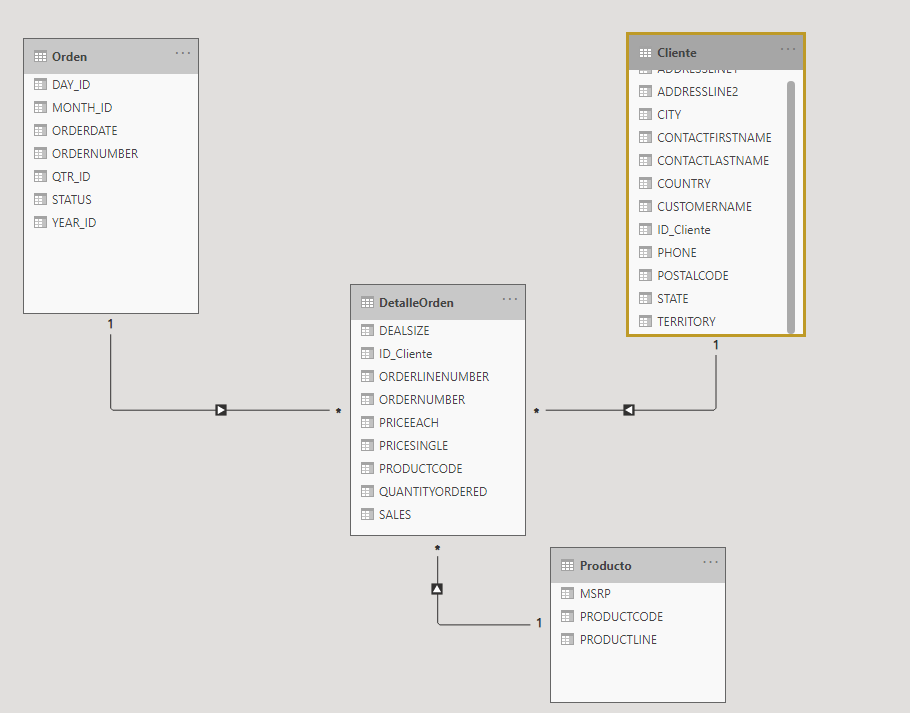
## Modelado de datos

Dentro de las variables se identifica cuatro tablas que son:

* Cliente: ID Cliente, Nombre del cliente, ciudad, país, dirección, teléfono, territorio, etc.
* Producto: Código del producto, precio base del producto.
* Orden: Número de orden de la venta, ID Cliente, Fecha de la venta, Año, Trimestre, Mes y Estado de la venta.
* Detalle de la venta: ID Cliente, Número de Orden de la venta, Código del producto, Cantidad ordenada, precio unitario del producto, venta del producto.

En el Gráfico 1se pude observar cómo quedaría modelada la información.

Gráfico 1Modelado de datos para las ventas de productos

Fuente: Elaboración propia

# Segmentación de Clientes

Para segmentar los clientes, ud podría utilizar la técnica RFM, la cual consiste en clasificar los clientes en función de su Reciente (Recency, Días transcurridos desde su última compra), Frecuencia (Frequency, Número de compras realizadas en un periodo determinado), Valor Monetario (Money, Valor de las compras totales realizadas en un periodo de tiempo) (IMF, 2017).

Está técnica consiste en calcular para cada cliente las variables anteriores. Para esté ejercicio, nuestros clientes son las empresas que compran y se sugiere tener en cuenta las ventas realizadas entre 6/01/2003 – 31/05/2005. Posteriormente, construya las variables RFM según la siguiente Tabla 1.

Tabla 1 Clasificación de las empresas según el RFM

|  |  |  |
| --- | --- | --- |
| *Recencia* | *Frecuencia* | *Valor Monetario* |
| 5, más reciente | 5, Más frecuente | 5, valor más alto |
| 4 | 4 | 4 |
| 3 | 3 | 3 |
| 2 | 2 | 2 |
| 1 más antiguo | 1, menos frecuente | 1, valor más bajo |

Fuente: Elaboración propia.

Teniendo en cuenta estos resultados, clasifican las empresas en 5 grupos (Tabla 5).

Tabla 5Clasificación de los grupos según los resultados del RFM

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Segmentos* | *R* | *F* | *M* | *Descripción* |
| Potenciales | 4,5 | 4,5 | 4,5 | Son los clientes que menos días tardan en comprar, los que más a menudo lo realizan y son los que más gastan. |
| Perdidos | 1,2,3 | 1,2,3 | 1,2,3 | Son clientes que compraron hace mucho tiempo, poca cantidad y poco gasto. |
| Derrochadores |  |  | 5 | Los que más gastan |
| Leales |  | 4,5 |  | Son los que más frecuente van a comprar |
| Nuevos | 4,5 |  |  | Clientes que han comprado hace poco pero no a menudo. |

Fuente: Elaboración propia

## Modelo de Sistema de Recomendación

En este ejercicio, se desea que ud proponga un modelo de sistema de recomendación. Tenga en cuenta que en cada orden se compraron varios productos y en diferentes cantidades. Por ejemplo, la siguiente imagen se puede evidencia que en la orden 10100 se compraron 4 productos en diferente cantidades (30, 50, 22, 49) y en la orden 10102 se compraron 2 productos también en diferentes cantidades (39, 41).

La idea es construir un modelo en cual, si un cliente compra un producto, se le pueda ofrecer otro producto. Identifique el producto que más se vendió y proponga por lo menos 3 combos. Lo mismo para el producto que menos se vendió.

