

A photograph of a young woman with long blonde hair, wearing a white t-shirt and jeans, standing in a supermarket aisle. She is looking down at a bunch of peaches she is holding in her hands. The supermarket has bright overhead lights and shelves filled with various products in the background.

**SQL**

**Alumno: Omar Miguel Cabbad**

**Profesor: Emilio D Aguero**

**Comisión: 81820**

# Tabla de contenido

1. Introducción .....	3
2. Descripción de la temática de los datos .....	4
3. Problemática.....	5
4. Objetivos....	6
5. Diagrama Entidad-Relación.....	7
6. Listado de campos por tablas.....	11

## 1. Introducción

En este mundo globalizado, las empresas deben estudiar cuidadosamente los caminos a seguir para consolidarse y diferenciarse de la competencia. Para ello, enfrentan desafíos como conocer el comportamiento de sus clientes, optimizar sus estrategias de ventas y gestionar eficientemente sus productos y sucursales. Un análisis adecuado de las ventas y la segmentación de clientes es clave para mejorar la rentabilidad y la satisfacción del consumidor.

En este proyecto se utilizará un dataset obtenido de Kaggle, una plataforma online ampliamente reconocida donde se comparten datasets, se realizan competencias de machine learning y se aprende sobre ciencia de datos. El dataset puede descargarse desde el siguiente enlace:  
<https://www.kaggle.com/datasets/faresashraf1001/supermarket-sales>

Este conjunto de datos contiene información detallada sobre ventas en supermercados, incluyendo datos de facturación, sucursal, ciudad, tipo de cliente, género, línea de productos, precios, cantidades, métodos de pago y márgenes de ganancia. Esto permitirá desarrollar análisis para identificar patrones de compra, evaluar el rendimiento de productos y clientes, y tomar decisiones estratégicas que contribuyan al logro de los objetivos empresariales.

## **2. Descripción de la temática de los datos**

La temática de la base de datos se centra en el funcionamiento de una cadena de supermercados y en cómo diversos factores influyen en su desempeño diario. La información incluye datos detallados sobre ventas, sucursales, ciudades, tipos de clientes, géneros, líneas de productos, precios, cantidades, métodos de pago y márgenes de ganancia.

El propósito de esta base de datos es organizar toda esta información de manera estructurada, permitiendo analizar el comportamiento de las ventas según diferentes variables, como el tipo de cliente, la ubicación, el producto o la forma de pago.

Contar con esta base de datos facilita el análisis del negocio desde múltiples perspectivas: identificar patrones de compra, evaluar el rendimiento de productos y sucursales, detectar oportunidades para mejorar la experiencia del cliente y optimizar las estrategias comerciales.

A través del uso de SQL, será posible realizar consultas que ayuden a comprender mejor la operación diaria, generar reportes relevantes para distintas áreas de la empresa y apoyar la toma de decisiones para aumentar la eficiencia y rentabilidad del supermercado.

### **3. Problemática**

Es importante implementar una base de datos para resolver problemáticas comunes en la gestión de ventas y operaciones en supermercados. En el mundo actual, las empresas necesitan manejar de forma eficiente la información relacionada con las ventas, clientes, productos y sucursales, optimizando los procesos desde la compra hasta la atención al cliente. Sin embargo, enfrentan dificultades para analizar el comportamiento de los consumidores y prever la demanda de productos.

Sin una base de datos centralizada y bien estructurada, pueden presentarse problemas que afectan el desempeño del negocio, tales como:

- Falta de datos consolidados sobre ventas por sucursal, tipo de cliente y línea de productos.
- Dificultad para evaluar el impacto de variables como métodos de pago, horarios o características del cliente.
- Limitada capacidad para tomar decisiones comerciales basadas en información precisa y actualizada.
- Necesidad de mejorar la formación del personal para el manejo adecuado de tecnologías y análisis de datos.

## **4. Objetivos**

### Objetivo General:

El objetivo general del proyecto es lograr una mejor gestión de ventas y desempeño de los supermercados mediante el análisis de datos históricos de transacciones y variables asociadas al cliente y al producto. Esto permitirá apoyar decisiones estratégicas, optimizar recursos y mejorar la planificación comercial y operativa de cada sucursal.

### Objetivos Específicos:

- Analizar los datos históricos de ventas por sucursal, línea de productos, tipo de cliente y método de pago.
- Evaluar el impacto de factores como horarios, días de la semana, promociones y características del cliente sobre las ventas.
- Identificar patrones de comportamiento de compra que permitan optimizar la planificación de productos y stock.
- Integrar información de distintas áreas (ventas, logística, contabilidad) para apoyar la toma de decisiones estratégicas.
- Proponer estrategias basadas en los análisis realizados que contribuyan a mejorar la eficiencia, rentabilidad y experiencia del cliente en los supermercados.

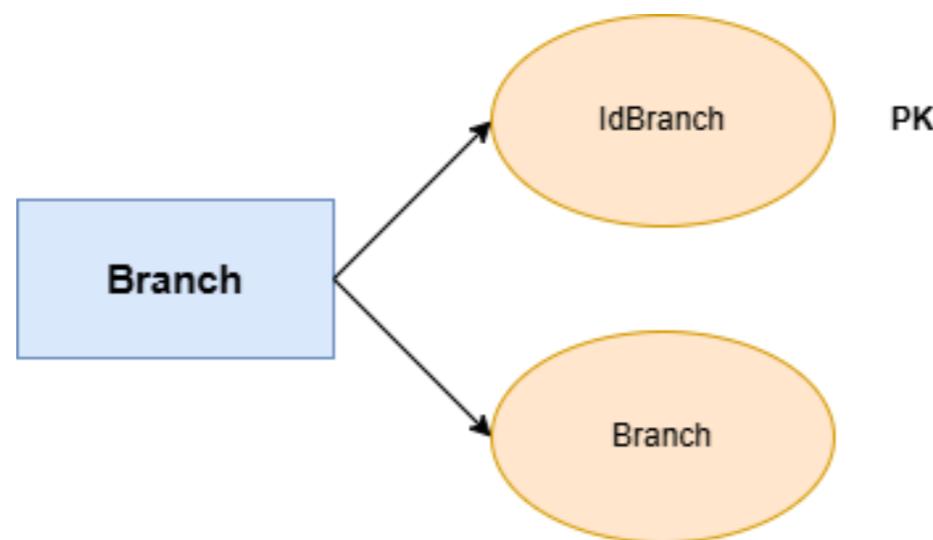
## 5. Diagrama Entidad-Relación

El modelo Entidad-Relación (ER) es un modelo de datos utilizado en el diseño de bases de datos. Representa entidades y sus relaciones en el mundo real.

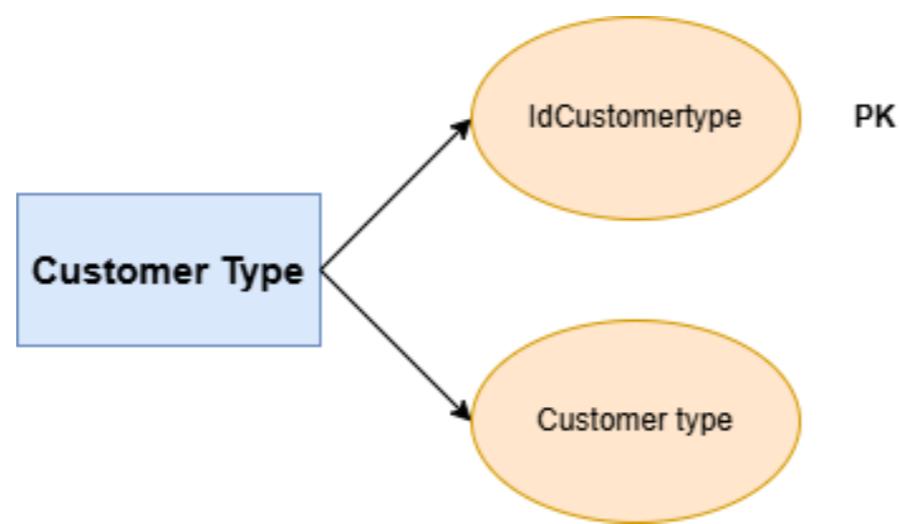
Para garantizar la correcta organización de los datos y permitir establecer relaciones entre las diferentes entidades del modelo, se incorporará un identificador único (ID) en cada tabla principal permitiendo una clave primaria en cada tabla (PK), asegurando un registro único, creación de claves foráneas (FK) y logrando una integridad diferencial.

Para nuestro proyecto tendremos lo siguiente:

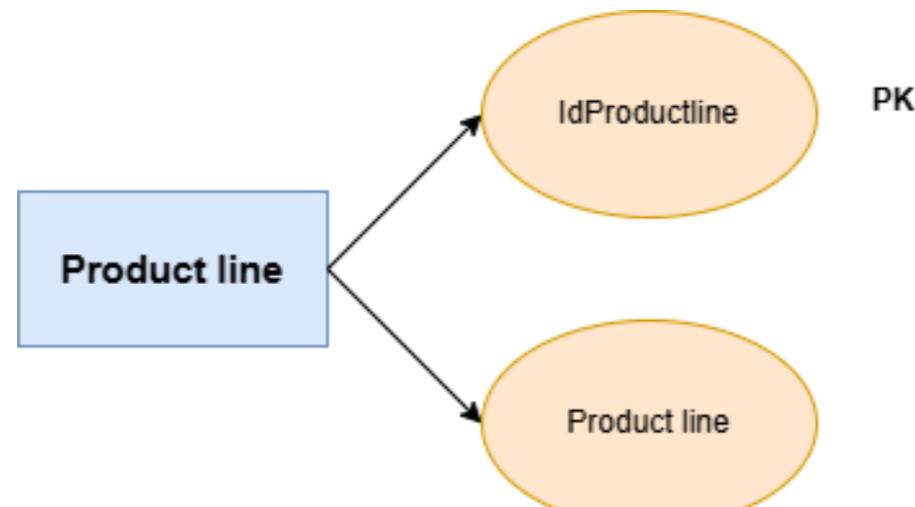
- La tabla **BRANCH** contiene la información sobre las distintas sucursales en donde se comercializa los productos. Incluye un ID para cada sucursal y el nombre de la sucursal.



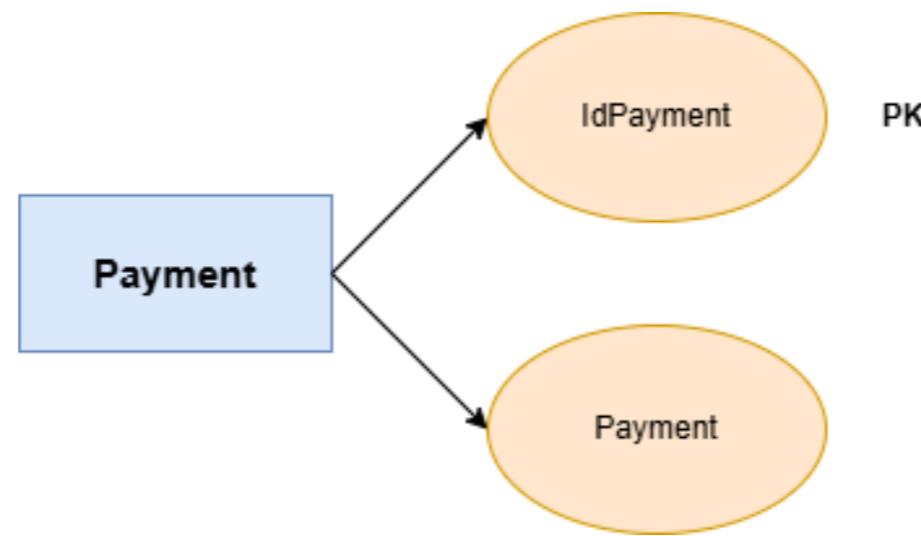
- La tabla **CUSTOMER TYPE** contiene la información sobre el tipo de cliente que compra el producto en el supermercado. Incluye un ID para cada tipo de cliente y el nombre del tipo de cliente.



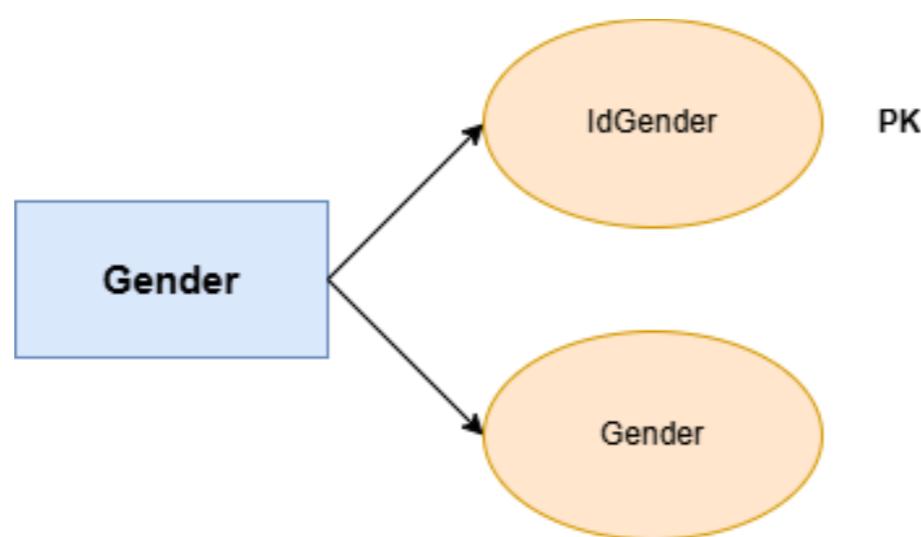
- La tabla **PRODUCT LINE** contiene la información sobre el tipo de cliente que compra el producto en el supermercado. Incluye un ID para cada tipo de cliente y el nombre del tipo de cliente.



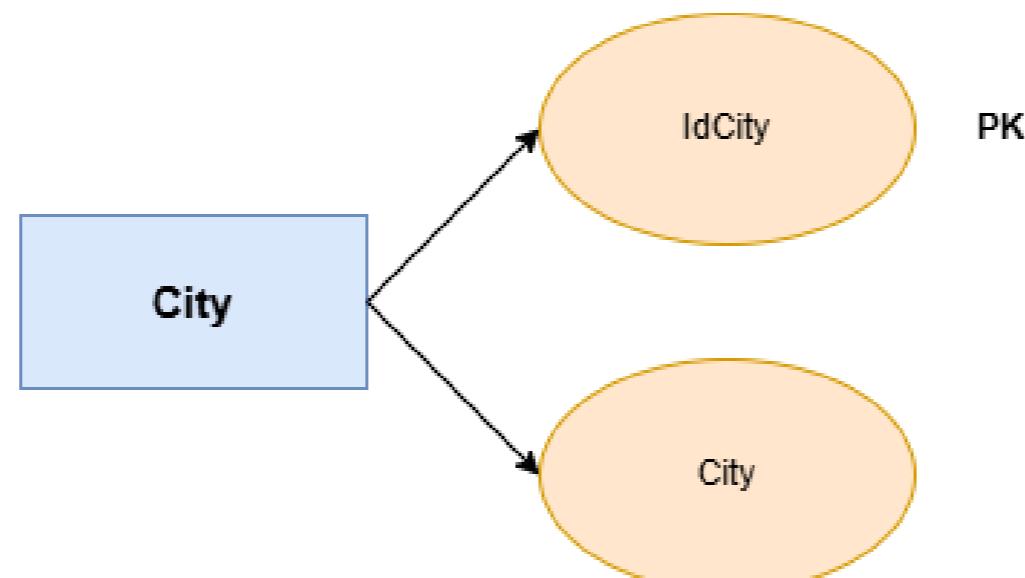
- La tabla **PAYMENT** contiene la información sobre la forma de pago en el cual se adquieren los productos del supermercado. Incluye un ID para cada método de pago y el nombre de dicho método.



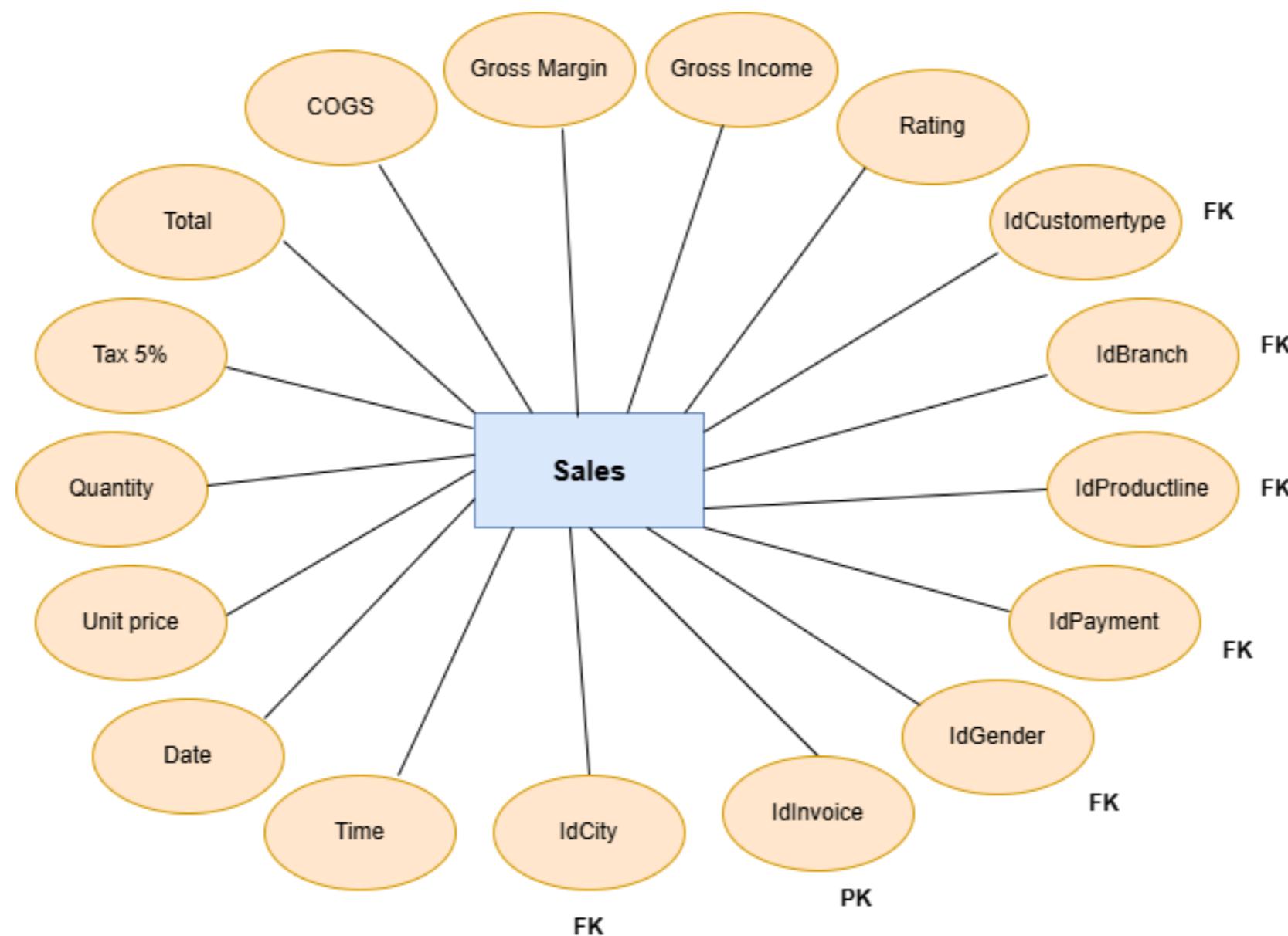
- La tabla **Gender** brinda información sobre los géneros que adquieren productos en el supermercado. Incluye un ID para el tipo de género y el nombre del tipo de género.



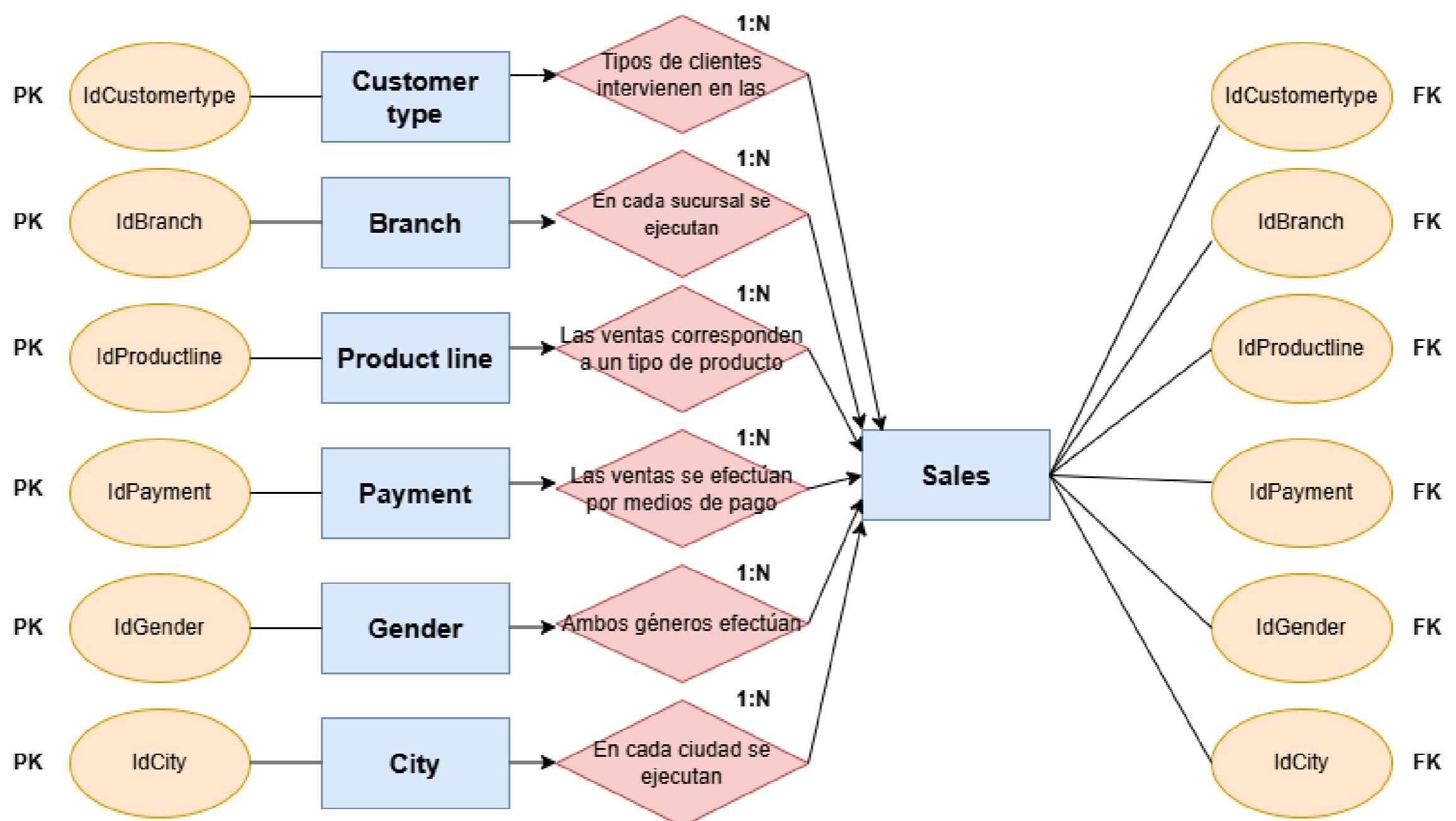
- La tabla **City** brinda información sobre las ciudades donde se comercializan los productos del supermercado. Incluye un ID para el tipo de ciudad y el nombre de la ciudad.



- La tabla **SALES** es la entidad central en la base de datos que registra cada transacción de venta realizada en las distintas sucursales. Contiene información relacionada con las dimensiones que describen las características de cada venta, tales como IDBranch, IDCity, entre otros. Ademas, incluye datos específicos de la transacción como InvoiceID, precio unitario, método de pago, cantidad vendida, entre otros,



- El modelo Entidad-Relación (ER) final quedaría de la siguiente manera:



## 6. Listado de campos por tablas

En el siguiente apartado, procederemos a proporcionar un exhaustivo desglose de estructura de cada tabla, incluyendo la enumeración de sus columnas, la especificación detallada de los tipos de datos asignados a cada una de ellas y, además, la clara identificación de los tipos de clave que se han implementado en dichas tablas. Este análisis minucioso de la disposición y características de los datos permitirá una comprensión más profunda y completa de la base de datos en cuestión, brindando una visión integral de su diseño y funcionamiento.

Tabla: BRANCH		
Campos	Tipo de datos	Claves
IdBranch	INT	PK
Branch	VARCHAR	

Tabla: CUSTOMER TYPE		
Campos	Tipo de datos	Claves
IdCustomertype	INT	PK
Customer type	VARCHAR	

Tabla: PRODUCT LINE		
Campos	Tipo de datos	Claves
IdProductline	INT	PK
Product Line	VARCHAR	

Tabla: PAYMENT		
Campos	Tipo de datos	Claves
IdPayment	INT	PK
Payment	VARCHAR	

Tabla: GENDER		
Campos	Tipo de datos	Claves
IdGender	INT	PK
Gender	VARCHAR	

Tabla: SALES		
Campos	Tipo de datos	Claves
IdBranch	INT	FK
IdCustomertype	INT	FK
IdProductline	INT	FK
IdPayment	INT	FK
IdGender	INT	FK
IdInvoice	INT	PK
IdCity	INT	FK
Time	INT	FK
Date	DATE	
Unit price	FLOAT	
Quantity	INT	
Tax 5%	FLOAT	
Total	FLOAT	
COGS	FLOAT	
Gross Margin	FLOAT	
Gross Income	FLOAT	
Rating	FLOAT	

