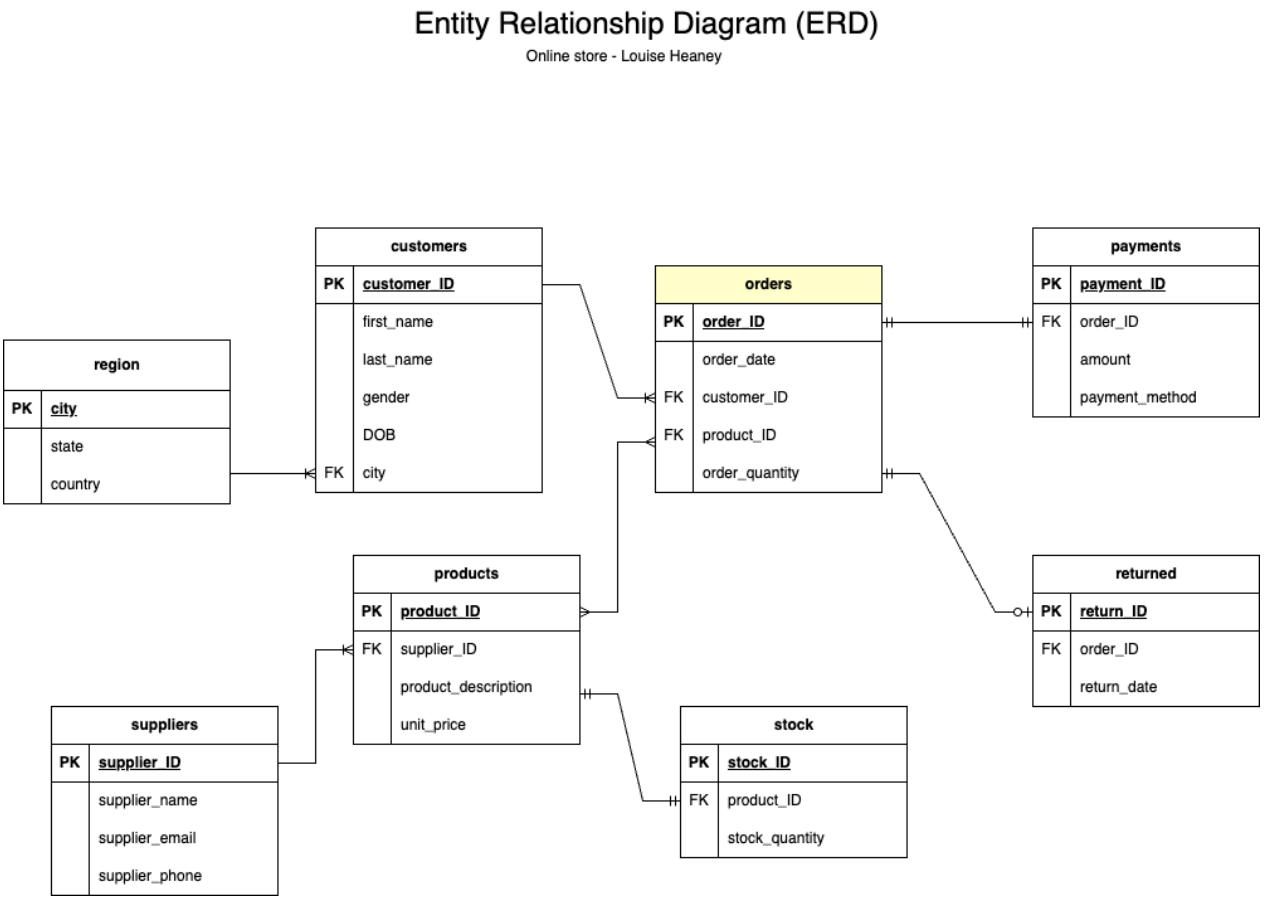
Database Design

The original dataset consisted of 1 large table:

Order ID	Order Date	Ship Date	Ship Mode	Customer ID	first_name	last_name	birthdate	gender	City	State	Countr
US-2023-151750	01/01/2023	05/01/2023	Standard Class	JM-15250	Niki	Archell	01/07/1967	Male	Huntsville	Texas	United
US-2023-144463	01/01/2023	05/01/2023	Standard Class	SC-20725	Ives	Phette	11/17/1988	Male	Los Angeles	California	United
US-2023-151750	01/01/2023	05/01/2023	Standard Class	JM-15250	Niki	Archell	01/07/1967	Male	Huntsville	Texas	United
US-2023-107503	01/01/2023	06/01/2023	Standard Class	GA-14725	Peta	Gosson	06/02/1975	Female	Lorain	Ohio	United
US-2023-151750	01/01/2023	05/01/2023	Standard Class	JM-15250	Niki	Archell	01/07/1967	Male	Huntsville	Texas	United
US-2023-151750	01/01/2023	05/01/2023	Standard Class	JM-15250	Niki	Archell	01/07/1967	Male	Huntsville	Texas	United
US-2023-151750	01/01/2023	05/01/2023	Standard Class	JM-15250	Niki	Archell	01/07/1967	Male	Huntsville	Texas	United
US-2023-151750	01/01/2023	05/01/2023	Standard Class	JM-15250	Niki	Archell	01/07/1967	Male	Huntsville	Texas	United
US-2023-154466	01/01/2023	02/01/2023	First Class	DP-13390	Hube	Simmig	8/20/1981	Male	Franklin	Wisconsin	United
US-2023-151750	01/01/2023	05/01/2023	Standard Class	JM-15250	Niki	Archell	01/07/1967	Male	Huntsville	Texas	United
US-2023-147207	02/01/2023	04/01/2023	Second Class	TS-21655	Ninon	Jacklings	2/26/1973	Female	El Paso	Texas	United
US-2023-147207	02/01/2023	04/01/2023	Second Class	TS-21655	Ninon	Jacklings	2/26/1973	Female	El Paso	Texas	United
US-2023-147207	02/01/2023	04/01/2023	Second Class	TS-21655	Ninon	Jacklings	2/26/1973	Female	El Paso	Texas	United

A flat file database like this is not ideal for a growing online business so, for efficiency reasons, I created a relational database which consisted of a main fact table supplemented by 7 dimension tables as illustrated below:



The main table (or entity), Orders, contains the unique transactional data for every order. Additional details such as Product details, customer details and supplier details have been separated out into their own entities, or normalised, to reduce repetition and data redundancy. However, all the data is still easily accessible and retrievable using their respective Primary/Foreign Keys.

These keys play a crucial role in the design and management of the database. The primary keys ensure that every row is uniquely identifiable, which prevents duplication of records and maintains data integrity. They also allow us to join tables together when used in combination with Foreign Keys.

The lines connecting the tables illustrate the relationships between these entities so for example there is a:

- **one to many relationship** between the Customer and Orders tables ie. one customer can place many orders
- **one to one relationship** between the Orders and Payments tables ie. an order can only have one payment method
- **many to many relationship** between the Orders and Products tables ie. many orders can include many products

The Datamart consists of the follwing views: