

# **Designing a Scalable Database System for Sol Café: A Case Study**

**Laura Margarita Sánchez Padilla**

**Professional Certificate – Data Science**

**Canadá  
2025**

## Contenido

<b>Abstract .....</b>	<b>3</b>
<b>1. Introduction.....</b>	<b>3</b>
<b>2. Mission and Objectives.....</b>	<b>4</b>
<b>3. Methodology .....</b>	<b>4</b>
<b>3.1 Entity identification through business análisis.....</b>	<b>4</b>
<b>3.2 Attribute specification based on user needs .....</b>	<b>4</b>
<b>3.3 Defining relationships and cardinality .....</b>	<b>5</b>
<b>3.4 Creating an Entity Relationship Diagram (ERD) .....</b>	<b>5</b>
<b>3.5 Enforcing data integrity constraints .....</b>	<b>5</b>
<b>3.6 Implementing database views for business insights.....</b>	<b>5</b>
<b>4. Database Design.....</b>	<b>5</b>
<b>4.1 Entities and Fields .....</b>	<b>6</b>
<b>4.2 Relationships .....</b>	<b>10</b>
<b>5. Data Integrity and Constraints .....</b>	<b>11</b>
<b>6. SQL Queries .....</b>	<b>11</b>
<b>Conclusion .....</b>	<b>16</b>

## Abstract

This case study presents the design and implementation of a relational database for *Sol Café*, a specialty coffee company offering ethically sourced beans and handcrafted drinks. The objective was to build a scalable and reliable database to support operations and enable data-driven decision-making. The project involved defining key entities, attributes, relationships, integrity constraints, and views to support real-world business scenarios including sales tracking, employee performance evaluation, and supplier inventory management.

## 1. Introduction

This project aims to explain the design of an SQL database intended to support decision-making for a coffee-selling company called **Sol Café**.

Sol Café is a coffee company committed to quality, sustainability, and community. In addition to selling roasted beans, it provides a curated menu of handcrafted drinks. With growing operations, the need for structured, accessible, and consistent data management has become critical.

The main interest of this type of business when collecting data is to obtain valuable information about their sales, which involve customers, employees, and the products being offered—since a sale cannot take place without any of these elements. That is why we sought to design a database that connects these agents in a coherent and structured way to ensure integrity and consistency.

This document presents the process of identifying tables and attributes, specifying constraints and data rules, defining relationships between entities and their respective ERD, as well as key aspects of database creation, including SQL views that are highly useful for business data analysis.

## **2. Mission and Objectives**

### **Mission**

To streamline operations and support growth by developing a reliable and scalable database that effectively manages essential business information and enables data-driven decision-making.

### **Objectives**

- Track and monitor sales activity effectively
- Evaluate employee performance
- Understand customer buying behavior
- Support accurate product and supplier tracking
- Maintain data consistency and referential integrity

## **3. Methodology**

To develop this database, we applied a methodology focused not only on accurately representing the business operations but also on enabling effective analysis of those operations. The following steps outline the process:

### **3.1 Entity identification through business análisis**

The first phase involved analyzing Sol Café's core business operations to identify the fundamental components of the system. Key processes such as sales, product management, supplier relationships, customer service, and shipping logistics were studied. As a result, seven essential entities were defined: Customers, Products, Sales, Employees, Suppliers, Store Locations, and Shipments. Each entity represents a dataset necessary for executing specific business functions.

### **3.2 Attribute specification based on user needs**

Once the entities were identified, we defined the specific attributes or fields required for each one. This step considered both business requirements and user needs. For instance, in the Employees table, fields such as name, role, phone number, and salary were included. In the Sales table, key fields like sale date, price, and references to other entities were determined. This specification allows for capturing precise, relevant data that is useful for daily operations and business analysis.

### **3.3 Defining relationships and cardinality**

With the entities and their attributes in place, the relationships between tables were analyzed to accurately represent how different parts of the business interact. For example:

- A customer can make multiple purchases (one-to-many).
- A supplier can provide several products.
- A sale is associated with one product and may be included in a shipment along with other sales.

We also defined the cardinality (one-to-one, one-to-many, many-to-many) of each relationship to reflect the business rules precisely. This step is critical to ensure the logical consistency of the model.

### **3.4 Creating an Entity Relationship Diagram (ERD)**

Using the defined entities, attributes, and relationships, an Entity-Relationship Diagram (ERD) was created. This diagram provides a visual representation of the database's logical model, showing tables, primary keys, foreign keys, and the direction of relationships. The ERD served as a roadmap for the physical implementation of the schema in a database management system (DBMS).

### **3.5 Enforcing data integrity constraints**

To ensure data accuracy and consistency, several integrity constraints were implemented. These rules ensure that the data entered into the system remains valid and consistent at all times.

### **3.6 Implementing database views for business insights**

Finally, SQL views were created to simplify access to consolidated and relevant information for decision-making. These pre-defined queries enable quick reporting without affecting system performance and provide users with a clear view of business performance.

## **4. Database Design**

The design of this database considered all aspects related to coffee sales. Initially, nine entities were proposed, but after filtering and analysis to avoid redundancy, the number was reduced to seven. For each entity, the most appropriate attributes were selected to ensure clarity and purpose.

#### 4.1 Entities and Fields

The final design includes the following core entities:

- Customers

This entity allows to track who is buying the products. Helps to give a personalized attention sending promotions or information related to the buying behavior.

Field Name	Data Type	Description	Constraints
CustomerID	Integer	Unique ID assigned to each customer	PRIMARY KEY, NOT NULL, UNIQUE
FirstName	Varchar (50)	Customer's first name	NOT NULL
LastName	Varchar (50)	Customer's last name	NOT NULL
Email	Varchar (100)	Customer's email	
Phone	Varchar (20)	Contact number	

- Employees

By managing employees data it is possible to track staff involved in the company. It helps to assign responsibilities and analyze performance.

Field Name	Data Type	Description	Key Type
EmployeeID	Integer	Unique ID for each employee	PRIMARY KEY, NOT NULL
FirstName	Varchar (50)	Employee's first name	NOT NULL
LastName	Varchar (50)	Employee's last name	NOT NULL
Role	Varchar (20)	Job title (e.g. barista, manager)	NOT NULL
Phone	Varchar (20)	Contact number	NOT NULL, UNIQUE
Salary	Decimal (10, 2)	Base Salary per month	NOT NULL, CHECK (SALARY>0)

- Products

The data about what the company sells is essential for tracking inventory and the sales trends of each producto.

Field Name	Data Type	Description	Key Type
ProductID	Integer	Unique ID for each product	PRIMARY KEY, NOT NULL, UNIQUE
Name	Varchar (20)	Product name (e.g. Colombian Roast)	NOT NULL
Type	Varchar (50)	Describes the type of bean or drink	NOT NULL
SupplierID	Integer	References de supplier providing the product	FOREIGN KEY, NOT NULL
Price	Decimal (10, 2)	Price per unit	NOT NULL, CHECK (PRICE>0)

- Suppliers

Stores information about Sol Café suppliers allows the company to manage the resource of their products. By knowing the origin of each product is easier to track quality and availability.

Field Name	Data Type	Description	Key Type
SuppliersID	Integer	Unique ID for each supplier	PRIMARY KEY, NOT NULL, UNIQUE
Name	Varchar (50)	Supplier's company or contact name	NOT NULL
Phone	Varchar (20)	Contact number	NOT NULL
Country	Varcha (20)	Country where the supplier is located	NOT NULL

- Store Locations

This table allows Sol Café to identify where the sale is happening. It is useful for analyzing performance by location, understanding regional preferences, and allocating resources appropriately..

Field Name	Data Type	Description	Key Type
StoreLocationID	Integer	Unique ID for each store	PRIMARY KEY, NOT NULL, UNIQUE
City	Varchar (25)	City where is located	NOT NULL
Country	Varchar (25)	Country of the store	NOT NULL
Address	Varcha (100)	Physical address	NOT NULL
Phone	Varchar (20)	Store contact number	NOT NULL

- Shipments

Manages delivery logistics. It shows which sales require shipping and their status making easier to track it.

Field Name	Data Type	Description	Key Type
ShipmentID	Integer	Unique ID for each shipment.	PRIMARY KEY, NOT NULL, UNIQUE
Date	Date	Date of dispatched	NOT NULL
Status	Varchar (20)	Shipped / delivered	NOT NULL

- Sales

The data in this core entity records every transaction which is important to provide data revenue, trends and performance metrics.

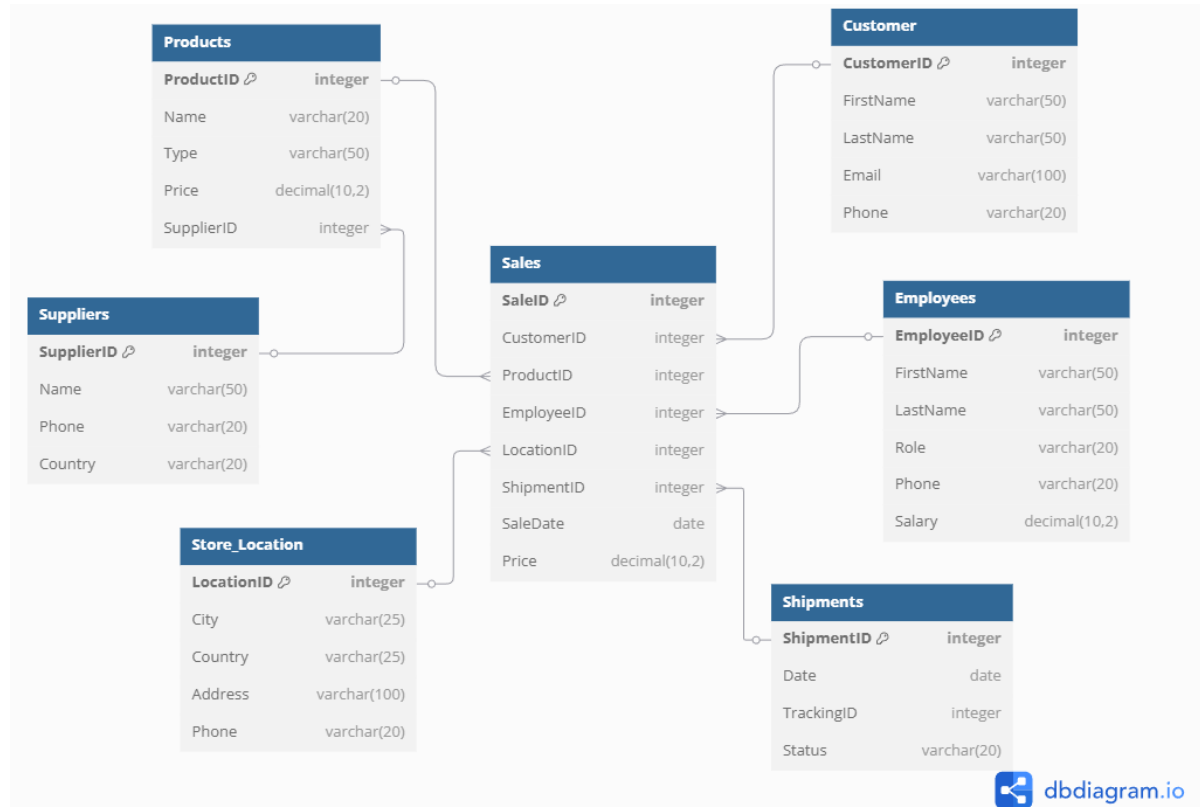
Field Name	Data Type	Description	Key Type
SaleID	Integer	Unique ID for each sale transaction	PRIMARY KEY, NOT NULL, UNIQUE



CustomerID	Integer	References the customer who made the purchase	FOREIGN KEY
ProductID	Integer	References the product sold	FOREIGN KEY
EmployeeID	Integer	References the employee who process the sale	FOREIGN KEY
LocationID	Integer	References the store where the sale was made	FOREIGN KEY
ShipmentID	Integer	Links a sale to its corresponding shipment	FOREIGN KEY
SaleDate	Date	Date the sale occurred	NOT NULL
Price	Decimal (10, 2)	Sale amount	NOT NULL, CHECK (PRICE >0)

## 4.2 Relationships

This is the Entity Relationship Diagram (ERD). It gives a simplify representation of our data structure.



Based on Sol Café's operations, the following relationships were defined:

- **Customer → Sales (1:M):** one customer can make many purchases, but each sale is made by only one customer.
- **Employee → Sales (1:M):** one employee can process many sales, but each sale is processed by one employee.
- **Store location → Sales (1:M):** one store location handles many sales, but each sale is made in only one store.
- **Product → Sales (1:M):** one product appears in many sales, but each sale has one product.
- **Supplier → Products (1:M):** one supplier provides many products, but each product is supplied by one provider.

- **Shipment → Sales (1:M):** one shipment delivers many sales, but there is only one shipment by sale.

## 5. Data Integrity and Constraints

The database includes several constraints to ensure consistency:

- **Primary keys** enforce uniqueness in all main entities and cannot be null. It prevents duplicate or incomplete records.
- **Foreign keys** maintain referential integrity. It links a column from a table to another table.
- **Check constraints** (e.g., Price > 0) prevent invalid data, mainly in prices and salaries.
- **Not null** fields ensure required information is always present for the essential fields.
- **Unique constraints** applied to fields like email and tracking IDs. Ensures that no more than a row in a table have the same value in the specified field.

## 6. SQL Queries

To support operational reporting and decision-making, some SQL queries were created:

- **Total Revenue:** Calculates overall income from sales which is useful to understand the financial performance of the company by calculating the sum of all sales made. This query helps to answer questions about sales goals and about the income earned in a specific period of time.

```
231  SELECT
232  |    SUM(Price) AS TotalRevenue
233  FROM Sales;
```

```
234
235  |
```

Results		Messages
	TotalRevenue	▼
1	336	

- **Sales by Each Employee:** Tracks performance of each staff member and how they contribute to business. It is important to take informed decisions about incentives and training needs.

```

235 SELECT
236     e.EmployeeID,
237     CONCAT(e.FirstName, ' ', e.LastName) AS EmployeeName,
238     COUNT(s.SaleID) AS TotalSales,
239     SUM(s.Price) AS TotalRevenue
240 FROM Sales s
241 JOIN employees e ON s.EmployeeID = e.EmployeeID
242 GROUP BY e.EmployeeID;

```

**Results**    Messages

	EmployeeID ▼	EmployeeName ▼	TotalSales ▼	TotalRevenue ▼
1	1	Sarah Johnson	11	126
2	2	Tom Clark	4	44
3	3	Nina Patel	5	60
4	4	Jake Miller	5	58
5	5	Lara Chen	5	48

- **Total Products Supplied by Each Supplier:** Monitors supplier contribution. It evaluates the role and contribution of each vendor in its supply chain. This information is essential for identifying key suppliers who provide a large portion of inventory and ensuring that the company maintains strong relationships with its most critical partners.

```

244 SELECT
245     s.SupplierID,
246     s.Name,
247     COUNT(p.ProductID) AS TotalProducts
248 FROM Suppliers s
249 JOIN Products p ON s.SupplierID = p.SupplierID
250 GROUP BY s.SupplierID;

```

**Results**    Messages

	SupplierID	Name	TotalProducts
1	1	Bean Brothers Co.	3
2	2	Colombian Gold	3
3	3	Sumatra Roast Ltd.	2
4	4	Kenya Highlands	1
5	5	Blue Mountain Roasters	1

- **View 1 Sales by location.**

Useful to understand which locations have better performance and which may need more attention.

```

165 CREATE VIEW SalesByLocation AS
166 SELECT
167     l.LocationID,
168     l.City AS StoreName,
169     COUNT(s.SaleID) AS TotalSales,
170     SUM(s.Price) AS Revenue
171 FROM
172     Store_Location l
173 JOIN Sales s ON l.LocationID = s.LocationID
174 GROUP BY l.LocationID;
175
176 SELECT * FROM SalesByLocation;
177

```

**Results**    Messages

	LocationID	StoreName	TotalSales	Revenue
1	1	New York	7	66
2	2	Toronto	11	124
3	3	Calgary	7	87
4	4	Edmonton	5	59

- **View 2 Most popular products**

This view helps to identify best selling products so Sol Cafecan manage inventory efficiently and know what products their customers prefer.

```

178 CREATE VIEW MostPopularProducts AS
179 SELECT
180     p.ProductID,
181     p.Name AS ProductName,
182     COUNT(s.SaleID) AS TimesSold,
183     SUM(s.Price) AS TotalRevenue
184 FROM
185     Products p
186 JOIN Sales s ON p.ProductID = s.ProductID
187 GROUP BY p.ProductID
188 ORDER BY TimesSold DESC;
189
190 Select * FROM MostPopularProducts;

```

**Results**    Messages

	ProductID	ProductName	TimesSold	TotalRevenue
1	1	Espresso Beans	7	81
2	8	Vanilla Latte	4	43
3	4	House Blend	3	32
4	9	Ethiopian Roast	3	34
5	10	Nitro Cold Brew	3	43
6	2	Arabica Blend	2	23
7	3	Cold Brew Bottle	2	20
8	5	Latte Can	2	11
9	6	Mocha Beans	2	26
10	7	Iced Americano	2	23

- **View 3 Most frequent customers**

Identifies loyal buyers who make repeat purchases, offering an opportunity to strengthen relationships through personalized rewards.

```

206 CREATE VIEW RepeatCustomers AS
207 SELECT
208     c.CustomerID,
209     CONCAT(c.FirstName, ' ', c.LastName) AS CustomerName,
210     COUNT(s.SaleID) AS PurchaseCount
211 FROM
212     Customer c
213 JOIN Sales s ON c.CustomerID = s.CustomerID
214 GROUP BY c.CustomerID
215 HAVING COUNT(s.SaleID) > 1;
216
217 Select * FROM RepeatCustomers;
218

```

Results Messages

	CustomerID	CustomerName	PurchaseCount
1	1	Alice Nguyen	4
2	2	Bob Smith	3
3	3	Carlos Diaz	5
4	4	Jack Neal	3
5	9	Richard Fields	3
6	10	Brian Newman	3
7	15	Jay Wise	3

- **View 4: Number of customers by City**

Reveals the geographic distribution of the customer base, which is essential for regional marketing, expansion planning, and tailoring services to local demand

```

219 CREATE VIEW CustomersByCity AS
220 SELECT
221     sl.City,
222     COUNT(DISTINCT c.CustomerID) AS TotalCustomers
223 FROM
224     Customer c
225 JOIN Sales s ON c.CustomerID = s.CustomerID
226 JOIN Store_Location sl ON s.LocationID = sl.LocationID
227 GROUP BY sl.City;
228
229 SELECT * FROM CustomersByCity;
230

```

Results Messages

	City	TotalCustomers
1	Calgary	5
2	Edmonton	4
3	New York	5
4	Toronto	9

These views help management monitor performance, adjust operations, and make informed decisions.

## **Conclusion**

The design and implementation of the Sol Café database address the operational and analytical needs of a growing coffee business. By organizing key entities such as customers, sales, products, employees, suppliers, store locations, and shipments, the database wants to ensure data consistency, traceability, and usability.

Through careful definition of relationships and constraints, the system maintains data integrity while supporting realistic business processes. This structure enables the company to track critical operations, from sourcing products to delivering orders and analyzing sales performance.

The creation of analytical queries such as total revenue, sales by employee, product supply by supplier, and customer behavior provides Sol Café with powerful tools to support data-driven decision-making. These insights help improve customer service, optimize inventory, assess employee productivity, and plan for strategic growth.