Database Design for Online Retail Store
Winter 2022

Abhimanyu Bhatnagar, 2020273
Madhava Krishna, 2020217
Ritika Nagar, 2020112
Atyant Sony, 2020039

April 7, 2022

Contents

Chapter 1

Defining the Problem

1.1 Objective

The goal is to design a database management system for an online retail store, similar to Big Bazaar, Flipkart and Amazon. Our project models on Grofers (now BlinkIt), a fast grocery marketplace for consumers to purchase day-to-day goods from.

1.2 Problem Statement

The ER diagrams and the relational schema that follow are based on the following (rudimentary) problem statement:

The online retail store serves many customers. The customers are required to hold an account on the platofrm to be able to purchase items. They can create an account by specifying their name, email address, phone number, and address. Customers add products to a shopping cart. They apply coupons on the shopping cart; the coupons have a coupon code and an associated discount percentage.
Customers order items by checking out the items on their shopping cart. The order is placed once the transaction is confirmed. The order is delivered to the customer only, they cannot have the order delivered to other addresses.
A product can belong to various categories and has specifications and a cost. Each product is obtained from a vendor which the store transacts with. After being purchased from the vendor, the products are stored in a warehouse.

Warehouse employees are responsible for packing and preparing orders. The readied order is then delivered to the customer by a delivery agent.
In case of any lapses with an order, the customer complains to support staff who create a complaint number against the order. They send out the details regarding the complaint to the customer.

1.3 Stakeholders

Notable stakeholders of this problem include:

1. Customer
2. Employees
3. Suppliers
4. Management of the company: board of executives, shareholders

Other stakeholders include manufacturers, warehousing service providers

1.4 Assumptions

1. The retail store would be operated in one country only. Therefore, there isn't an option to change the country.
2. There will be a shopping cart associated with the customer's account. They won't be able to add item to cart without being logged in.
3. The delivery would be taken care of by the company itself; items will be shipped from one warehouse only (Similar to Blinkit, erstwhile Grofers).
4. Coupons would be applied on the order, not the cart. Coupon would be applied during the checkout process.
5. Coupons would be applied using the coupon code, which is a unique alphanumeric value. A coupon cannot be reused.
6. Employees will be divided into delivery partners (responsible for delivering the order), warehouse workers (tasked with preparation of orders) and service employees (responsible for conflict resolution).

7. Vendors have only one account number and will be based only in India.
8. Employees have only one email address.
9. There won't be any wishlisting features, nor any saved-items feature like the ones offered by Amazon.
10. Employees' performance would be graded on a scale from 1 to 10 (for ease of data entry) with decimal values being permissible.
11. Discounts will be applied before taxes are calculated. Taxes will be calculated on the total amount after applying discount.

Chapter 2

ER Diagram

2.1 Entities

1. Customer
    (a) Customer ID: Primary key
    (b) Phone Number: Multivalued
    (c) Name: composite
        i. First Name
        ii. Last Name
    (d) Address: composite
        i. House Number
        ii. Locality
        iii. City
        iv. State
        v. Pin Code
    (e) Email Address
    (f) Password
2. Product
    (a) Product ID: Primary key
    (b) Specifications

i. Expiry date
ii. Manufacture date
iii. Country of manufacture
iv. Weight
v. Dimensions

(c) Price
(d) Category
(e) Discount Percentage
(f) GST
(g) Rating : Multivalued
(h) Photos: Multivalued

3. Orders

(a) Order ID
(b) Cost of products
(c) Taxes
(d) Discount Percentage

4. Vendor

(a) Vendor ID
(b) Address
i. Plot number
ii. City
iii. State
iv. Pin code
(c) Name
(d) Phone number

5. Coupon

(a) Coupon Code: Primary key
(b) Discount Percentage

(c) Validity

6. Warehouse
   (a) Warehouse ID
   (b) Address: composite
      i. Plot number
      ii. City
      iii. State
      iv. Pin Code
   (c) Telephone Number: multivalued
7. Employee:
   (a) Employee ID: Primary key
   (b) Date of Joining
   (c) Position
   (d) Department
   (e) Email Address
   (f) Name: composite
      i. First Name
      ii. Last Name
   (g) Residential Address: composite
      i. House number
      ii. Locality
      iii. City
      iv. State
      v. Pin Code
   (h) Date of Joining
   (i) Performance
   (j) Salary
   (k) Gender
   (l) Date of Birth

(m) Age : derived

Employees can be specialised into:

- Delivery Partner:
  (a) Vehicle ID
  (b) Vehicle Type
- Warehouse Worker: no additional attributes
- Service Employee: no additional attributes

8. Shopping Cart (Weak entity)
   (a) Customer ID,Product ID: Primary Key
   (b) Customer ID: Foreign Key (references Customer)
   (c) Product ID: Foreign Key (references Product)
   (d) Quantity
   (e) Total cost: Derived attribute

2.2 Relations

1. Supplies: between vendor and product
   (a) Vendor ID, Product ID: Primary Key
   (b) Vendor ID: Foreign Key (references Vendor)
   (c) Product ID: Foreign Key (references Product)
   (d) Quantity: attribute
2. Includes: between product and order to show which order contained what products
   (a) Order ID, Product ID: Primary Key
   (b) Order ID: Foreign Key (references Order)
   (c) Product ID: Foreign Key (references Product)
   (d) Quantity: attribute

3. Stores: between warehouse and product to show which warehouse contains what products.
    (a) Warehouse ID, Product ID: Primary Key
    (b) Warehouse ID: Foreign Key (references Warehouse)
    (c) Product ID: Foreign Key (references Product)
    (d) Stocks: attribute
4. Delivery: Quarternary relation between warehouse, order, delivery partner and customer to represent delivery.
    (a) Order ID: Primary Key
    (b) Order ID: Foreign Key (references Order)
    (c) Employee ID: Foreign Key (references Employee)
    (d) Customer ID: Foreign Key (references Customer)
    (e) Warehouse ID: Foreign Key (references Warehouse)
5. Transaction: Relation between customer and their cart, order and coupon code.
    (a) Order ID: Primary Key
    (b) Customer ID: Foreign Key (references Customer)
    (c) Payment Method
    (d) Transaction Status
    (e) Transaction Time
    (f) Coupon Code: Foreign Key (references Coupon)
6. Complains: between service employees (specialisation), customer and order to indicate a dispute.
    (a) Complaint Number: Primary Key
    (b) Customer ID: Foreign Key (references Customer)
    (c) Order ID: Foreign Key (references Order)
    (d) Service Employee ID: Foreign Key (references Employee)
7. Works in: between Warehouse Worker and warehouse.

(a) Employee ID, Warehouse ID: Primary key

8. Has: relation between customer and shopping cart (weak entity).

   (a) Customer ID: Primary key

9. Contains: relation between shopping cart and product to indicate that shopping cart contains the item.

   (a) Quantity

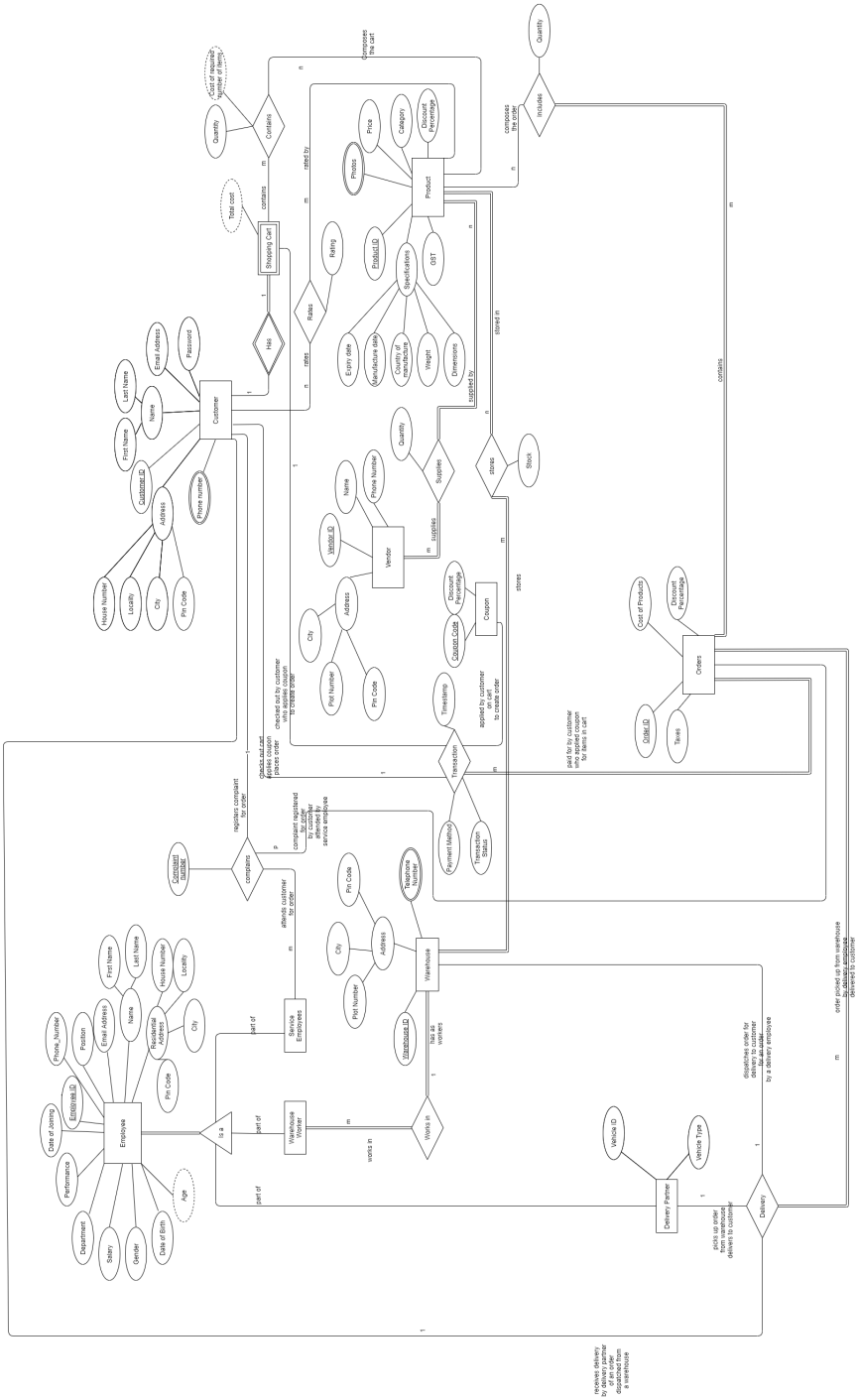
   (b) Cost of items in the cart: Derived quantity.

10. Rates: between customer and product for indicating rating.

    (a) Customer ID, Product ID: Primary key

    (b) Rating: attribute

2.3 Diagram


Customer
Customer ID
Name
First name
Last name
Email Address
Password
Address
House Number
Locality
City
Pin Code
Phone number
Shopping Cart
Total cost
Contains
Quantity
Count of items
Product
Product ID
Price
Category
Discount Percentage
Photos
Specifications
GST
Rates
Has
rates
rated by
Expiry date
Manufacture date
Country of manufacture
Weight
Dimensions
composes the cart
composes the order
includes
stored in
supplied by
Supplies
stores
Shop
Vendor
Vendor ID
Name
Phone Number
Address
City
Plot Number
Pin Code
supplies
Coupon
Coupon Code
Discount Percentage
checked out by customer who applies coupon to create order
applied by customer to get discount to create order
Timestamp
Transaction
Payment Method
Transaction total
paid for by customer who applied coupon to check out cart
Orders
Order ID
Taxes
Cost of Products
Discount Percentage
contains
Complaint raised
complains
registers complaint for order
complaint registered by customer checked by service employee
attends customer for order
Employee
Employee ID
Name
First Name
Last Name
Email Address
Phone_Number
Position
Residence Address
House Number
Locality
City
Pin Code
Date of Joining
Performance
Department
Salary
Gender
Date of Birth
Age
Is a
part of
Service Employee
Warehouse Worker
Warehouse
Warehouse ID
Telephone Number
has an address
Works in
works in
Delivery Partner
Vehicle ID
Vehicle Type
Delivery
dispatches order for delivery to customer by a delivery employee
picks up order from warehouse delivers to customer
order picked up from warehouse delivered to customer
receives delivery by delivery partner after order dispatched from a warehouse

Chapter 3

Relational Schema

The ER diagram was reduced by noting multiplicities and the following tables for the relational schema resulted.

1. Customer(customer ID, First_name, Last_name, House_number, Locality, City, Pincode, Email_Address, Password)
2. Product(product ID, Price, Category, Discount Percentage, GST Percentage)
3. Product Rating(Product ID, Customer ID, Rating)
4. Product Photo(Product ID, Product Photo)
5. Vendor(Vendor ID, First Name, Last Name, Plot Number, City, Pincode)
6. Vendor Phone(Vendor ID, Phone Number)
7. Warehouse(Warehouse ID, Plot Number, City, Pincode)
8. Warehouse Phone(Warehouse ID, Phone Number)
9. Employee(Employee ID, First name, Last name, Age, Salary, Gender, Department, Performance, Position, Date of Joining, Date of Birth, Email Address, House Number, Locality, City, Pincode, Phone Number)
10. Delivery Partner(Employee ID, Vehicle ID, Vehicle Type)

11. Warehouse Worker(Employee ID)

12. Service Employee (Employee ID)

13. Orders(Order ID, Total Price, Taxes, Total Discount Percentage)

14. Coupon(Coupon Code, Discount Percentage)

15. Transaction(Order ID, Payment Method, Transaction Status ,Transaction Time, Customer ID, Coupon code)
    Here, Customer ID is referenced from Customer table and coupon code from Coupon table. Coupon code can be null, customer id cannot.

16. Delivery (Order ID, Employee ID, Customer ID, Warehouse ID)
    Here Employee ID is taken from the Employee table if it belongs to Delivery Partner, Customer ID from Customer table, Warehouse ID from Warehouse table.

17. Stores(Warehouse ID, Product ID, Quantity)

18. Supplies(Vendor ID, Product ID, Quantity)

19. Shopping Cart(Customer ID, Product ID, Quantity)

20. Complains(Complaint Number, Customer ID, Order ID, Service Employee ID)
    Here, Customer ID is taken from Customer table, Order ID from Orders table, Service Employee ID from Service Employees table.

21. Order Products(Order ID, Product ID, Quantity)

Note: Employee has both fields of age and DOB. This is a redundancy on our part as the data entries were designed with age as an attribute. This may be eliminated in the final version depending on whether there is need or not (number of employees should be significantly lower than the number of customers).
The SQL queries for creating the database, populating the database and the queries required are submitted under the names: 'database_creation.sql', 'database_population.sql' and 'queries.sql' in the Database folder.

Chapter 4

Project Development Procedure and Deliverables

4.1 ER Formulation

Before commencing with the ER diagram formulation, many brainstorming sessions were held to determine which all aspects the online retail store would target and to what degree. The result of those sessions is the *problem statement* and the list of assumptions. The problem statement is a rudimentary sketch of what all scenarios the store should be equipped to handle and what to expect from the usage scenarios of the database management system. The ER diagram and relational schema have been added in the previous chapters and the diagram should be available in the ER_Diagram folder.

4.2 Weak Entity

The shopping cart was determined to be a weak entity since it does not have any identifying attribute other than belonging to a specific customer. Hence, on careful consideration of all its attributes and the lack of a distinctive attribute, it was decided to be kept as a weak entity.

4.3 Entity Relationship Participation and Types

This was another topic on which multiple meetings had been conducted, specifically on the nature of the ternary/quarternary relations and how to decompose them. Ultimately, Delivery and Transaction were kept because of their utility and ability to define the whole scenario without conflicts with other entities and relations. The participation types (total and partial), relationship roles and cardinality constraints have been mentioned in the ER diagram. Detailed constraints have been mentioned in the SQL database creation file.

4.4 Ternary Relation

A ternary relation: complains was identified which involves an order, the customer who placed that order and one of the service employees. The cardinality is many on the order's and service employee's side and unary on the customer's side. It was identified in the following way: "A customer can complain to many service employees about many orders."

4.5 Relational Schemas

They have been included in one of the earlier chapters.

4.6 Sufficient and Valid Constraints in DDL

The DDL file has been enclosed in the Database folder under the name 'database_creation.sql'.

4.7 Data Entry

Data was generated using Mockaroo and using python scripts for the foreign key tables. The primary keys were chosen in increasing order (similar to AUTO_INCREMENT) and most of the tables (except warehouse and complains, which contain 5 and 10 rows respectively), have 50 to over a hundred entries.

4.8 Queries

Queries have been included in the Database folder under the file name 'queries.sql', which should be run after running 'database_creation.sql' and 'data_population.sql' respectively.

Chapter 5

Contributions of Team Members

Collaboration was done by the means of Microsoft Visual Studio Code's Live Share feature which allowed for real time simultaneous editing of multiple files by many participants. Communication occurred through Google Meet links. Following are the contribution of each team member in the development and ideation of the project till the mid-evaluation. Many of the responsibilities overlapped and all tasks were completed with contributions from all.

1. Abhimanyu Bhatnagar:
    (a) Ideation of the ER diagram
    (b) Drawing the ER diagram
    (c) Reduction to Relational Schema
    (d) DDL : creating tables
    (e) Creating test data
    (f) Part of the SQL queries
    (g) Assisted with documentation
2. Atyant Sony:
    (a) Ideation of the ER diagram
    (b) Drawing the ER diagram
    (c) Reduction to Relational Schema

(d) DDL : creating tables

(e) Creating test data

(f) Part of the SQL queries

(g) Documentation and latex formatting.

3. Ritika Nagar:

(a) Ideation of the ER diagram

(b) Reduction to Relational Schema

(c) Creating test data

(d) SQL queries

4. Madhava Krishna:

(a) Determining requirements and setting a problem statement.

(b) Drawing the ER diagram

(c) Reduction to Relational Schema

(d) Part of the DDL

(e) Creating part of the test data

(f) SQL query

(g) Documentation and latex formatting.