ElectroBase Management System

EBMS - *an Online Electronics Retail Store*, built as a course project for
CSE202: Fundamentals of Database Management Systems

Project Report and Documentation

]Divyajeet Singh (2021529) Mehar Khurana (2021541)

March 17, 2023

Deadline 1

Defining Project Scope and Requirements

January 25, 2023

Project Scope

Electronics has always been a booming industry. With the advent of the internet, the industry has seen a massive shift towards online retail. However, it is difficult to keep track of the technical requirements of a store. It is difficult to keep all stakeholders in the loop and keep them involved and updated.

This is where we come in with **EBMS**, i.e. the **ElectroBase Management System**. EBMS is an online retail platform for electronics. It aims to provide a common platform for suppliers, store managers, customers, and delivery agents.

- It is an easy solution for the **customers**, as it aims to provide a diverse catalogue of products to customers. The customers get to choose from a wide range of categories, make changes to their cart, and make secure payments with a method of their choice.
- The **suppliers** get to keep track of their products and change their description, price, etc. They can also keep track of their sales statistics and make changes to their products as and when required.
- The database managers (**admins**) get assisted in monitoring the transactions and managing the inventory. Based on their requirements, they can add deals or combos on the available products or remove categories from their store.
- EBMS provides a platform for **delivery agents** to keep track of all orders that have been assigned to them. They can set their activity/inactivity status and view the feedback given to them.

The primary focus of the project is to design an efficient backend. We aim to create a system that is smooth and easy to use for the customers and easy to manage for the suppliers. The system should support efficient searching through the catalogue and should be able to handle a large number of transactions.

The backend will be built using MySQL, along with Python and Django, and will be hosted on a server. The frontend will be built using HTML, CSS, and ReactJS. By the end of the semester, we plan to host this project on a public server and make it accessible publicly.

TL; DR

The aim of this project is to bring to life an integrated online retail store for electronics. The project will bring all stakeholders on a common platform and will ensure a smooth and easy-to-use experience for the customers.

Technical Requirements

Tech Stack

We plan for EBMS to be a full-stack project with a backend and a frontend. According to the requirements, we plan to use the following tools and technologies:

- **MySQL Database**
- **Python-3**
- **Django Framework**
- **HTML**
- **CSS**
- **ReactJS**

Entities, Relations, & Constraints

Entities

The following entities are identified and will be used in the project:

1. **Admin:** Admins are the store managers. They are the stakeholders responsible for managing the inventory and maintaining the store (database).
2. **Customers:** Customers are the primary stakeholders interacting with the system. They must create an account and log in to use the app.
3. **Suppliers:** Suppliers are responsible for supplying products to the store. Different suppliers are allowed to supply the same products to the store.
4. **Delivery Agents:** A delivery agent will be assigned to each order. They will be responsible for delivering orders.
5. **Products:** As the name suggests, these are the products supplied to the store by suppliers. Products are weak entities since several sellers can sell the same product, ie, we can have several records for the same productID.
6. **Orders:** An order entity is used to keep track of the orders placed by different customers. Each order also has one associated delivery agent.
7. **Product Reviews:** Product reviews are the feedback given by customers on products. A customer can only review a product they have previously purchased.
8. **Delivery Agent Reviews:** Delivery agent reviews are the feedback given by customers to delivery agents. A customer can only post a review on a delivery agent they have received an order from.

9. **Wallet:** A wallet is a belonging of a customer that stores attributes like the current balance of the customer, their UPI-ID etc. Users' wallets are hidden from the admin's view. Wallets are weak entities since several people can be using the same UPI-ID, and we require the customerID to uniquely identify a wallet.

10. **Description:** The description weak entity is used to store the description of different products. This is used to avoid redundancy of data storage (in the case where multiple suppliers are selling the same products with the same descriptions).

Relationships & Cardinality Constraints

To effectively manage the database, we will be using the following (non-exhaustive list of) relationships among the data:

1. ***Cart*: Cart** is a relationship between **Customer** and **Product**. One customer can add multiple products to the cart (and choose a specific supplier). The same products may be added to more than one cart. When a customer checks out, the cart is used to generate an order entity, and a delivery agent is assigned to it.

2. ***Sells*: Supplier** *sells* **Product**. This many-to-many relationship is used to keep track of what products one supplier sells. Each supplier can supply multiple products. The same product may be sold by different suppliers.

3. ***Sold*: Supplier** has *sold* **Product**. This many-to-many relationship covers the products sold by a supplier and will be used to generate their sales statistics.

4. ***Delivered*: Delivery Agent** has *delivered* to **Customer**. This is a ternary relationship involving **Delivery Agent Review** as well. There can be, at most, one (editable) review from one customer for a delivery agent.

5. ***Purchased*: Customer** has *purchased* **Product**. This is a ternary relationship involving **Product Review** as well. There can be, at most, one (editable) review from one customer for a product.

6. ***Describes*: Description** *describes* **Product**. This is a one-to-one relationship.

7. ***Belongs To*: Wallet** *belongs to* **Customer**. This is a one-to-one relationship.

8. ***Consists Of*: Order** *consists* **Products**. This is a one-to-many relationship, as one order may consist of multiple products. The product quantities are moved from the cart to the order on checkout.

9. ***Manages*: Admin** *manages* the app. This relationship is used to hide entities like wallets from the admin's view.

Most cardinality constraints have been mentioned above in the description of relationships. Some other constraints are as follows:

1. **Existential:**
   (a) **Product - Supplier:** There can be no product without a supplier.
   (b) **Order - Customer:** An order cannot exist without a customer.
   (c) **Order - Delivery Agent:** An order cannot exist without a delivery agent.
   (d) **Product Review - Customer *has purchased* Product:** To ensure that a customer can review only the products they have purchased.
   (e) **Delivery Agent Review - Delivery Agent *delivered to* Customer:** to ensure that a customer can review only the delivery agents they have received orders from.
2. **One-to-One: Delivery Agent - Order:** One delivery agent can be assigned at most one order at a time. We may remove this constraint later and implement an algorithm to find the best agent to assign an order to, based on the current delivery addresses, with a cap on the number of orders per agent.
3. **One-to-Many: Customer - Phone Number** One customer can have multiple phone numbers. It will be implemented as a multi-valued attribute of Customer.

Access Constraints

Since some data must remain private while other data must be accessible to all, the following access-control constraints will be implemented:

1. **Admin:** All data except for Customer passwords and Wallets
2. **Customer:** Personal records, past and current Orders, and all Products and Reviews
3. **Supplier:** Their Product catalogue, personal records, sales statistics, and Customer Reviews
4. **Delivery Agent:** Personal records, current Orders, and Reviews from past Orders

Functional Requirements

All stakeholders except the **Admins** will need to create an account and log in. The following (non-exhaustive list of) features are some of the functional requirements of the project:

1. As a **Customer:**

- Add balance to wallet
- Browse and search/sort/filter for products
- Manage (add/remove) items in their cart
- Place an order (checkout cart)
- Confirm/Authenticate transaction
- View and search for previous orders

2. As an **Admin:**
    - Add/Delete categories
    - Delete products
    - Add/Modify discounts
    - Create deals/combos
    - View transaction history
    - Appoint other Admins
3. As a **Supplier:**
    - View products currently on sale
    - Add/Discontinue a product
    - Change price of a product
    - Change quantity of a product
    - View sales statistics
4. As a **Delivery Agent:**
    - Confirm a delivery
    - View the address of the order
    - View the ETA of the order
    - View reviews
    - Choose current activity status

Deadline 2

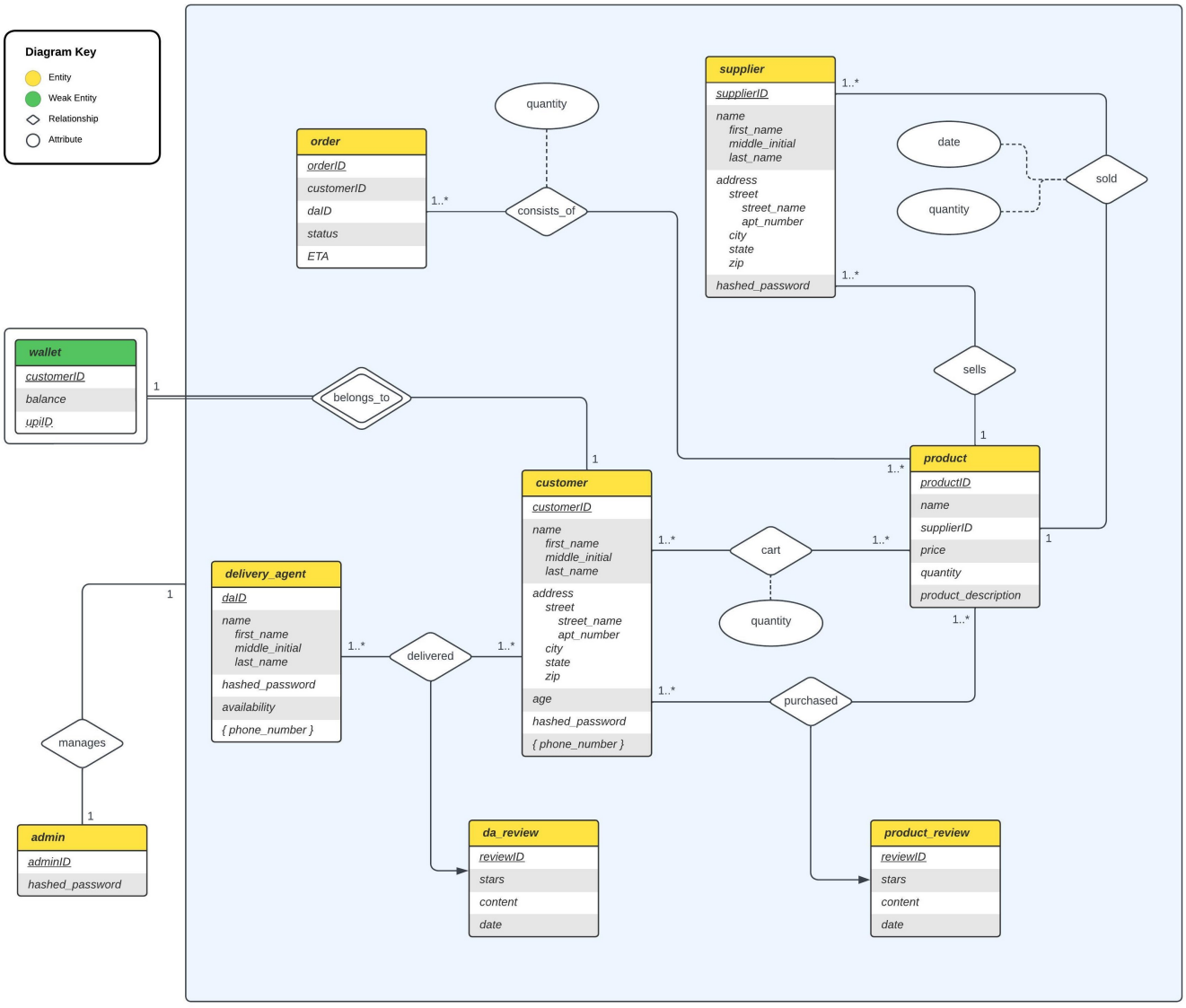
Designing the ER Model and converting it to a Relational Model

February 3, 2023

Entity-Relationship Model

Entity-Relationship (ER) Models are used to plan how different entities in a project interact with each other.

Our ER Model captures the nature of the relationships and entities planned to be used in the project. The ER Model is designed in accordance with the assumptions and constraints as mentioned in the document above. Hence, we plan to build our system on the basis of the following Entity-Relationship Model:



Ternary Relationships

The following ternary relationships have been identified:

1. **Customer - Product - Product Review:** A customer can review multiple products, and a product can be reviewed by multiple customers. A customer can give at most one review per product. This ternary relationship will be decomposed into the following binary relationships at the time of implentation:

   (a) **Customer - Product:** (Many-to-Many) To keep track of which customers have purchased which products.

   (b) **Product - Product Review:** (One-to-Many) To keep track of all reviews given to a product.

   (c) **Customer - Product Review:** (One-to-Many) To keep track of all reviews given by a customer. This relationship may not be needed and may be removed in the future.

2. **Delivery Agent - Customer - Delivery Agent Review:** A customer can review multiple delivery agents, and a delivery agent can be reviewed by multiple customers. A customer can give at most one review per delivery agent. This ternary relationship will be decomposed into the following binary relationships at the time of implentation:

   (a) **Customer - Delivery Agent:** (Many-to-Many) To keep track of which customers have received orders from which delivery agents.

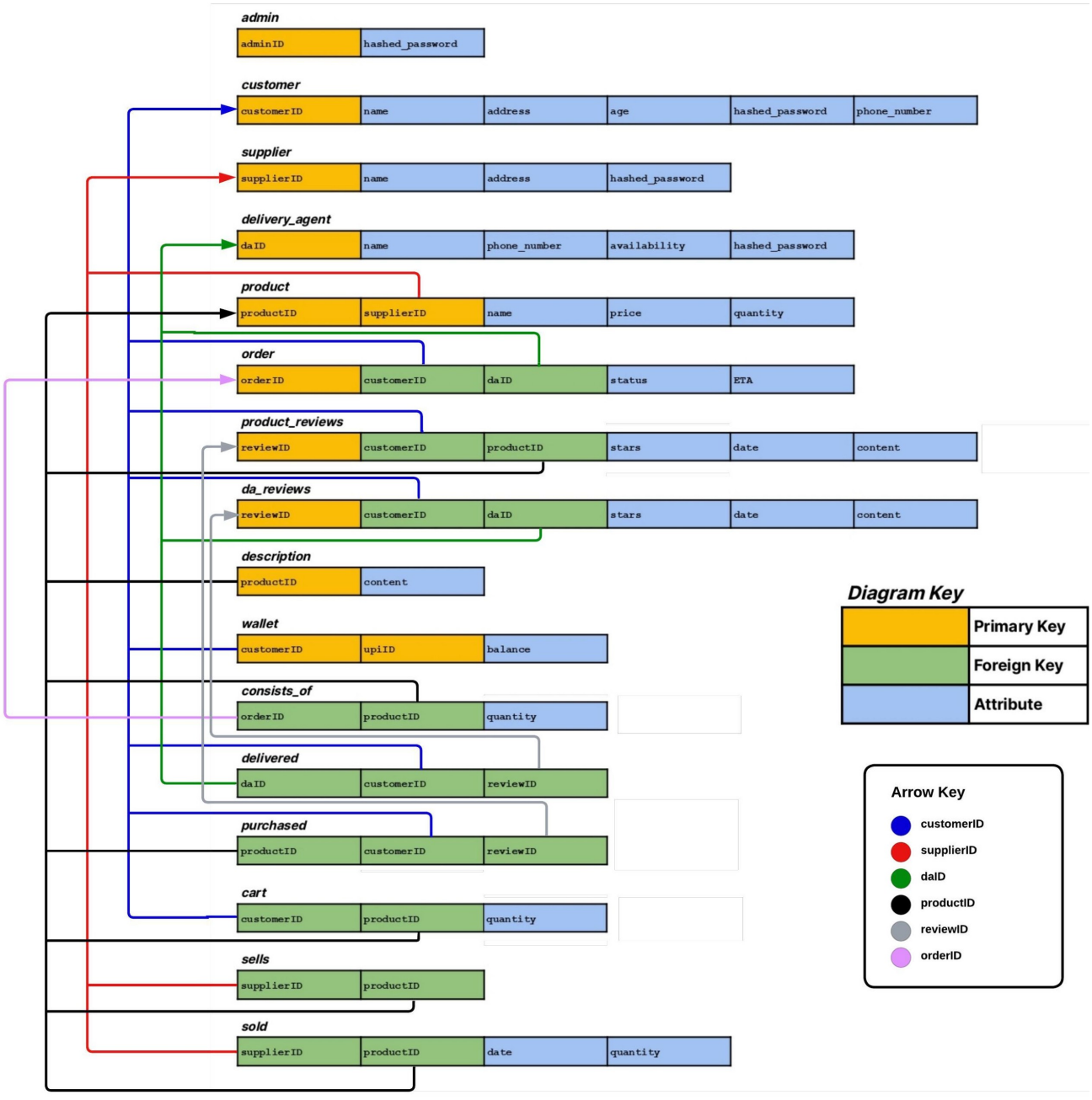
   (b) **Delivery Agent - Delivery Agent Review:** (One-to-Many) To keep track of all reviews given to a delivery agent.

   (c) **Customer - Delivery Agent Review:** (One-to-Many) To keep track of all reviews given by a customer. This relationship may not be needed and may be removed in the future.

Relational Model

Relationship Models are used to represent how data will be stored in the database, along with the attributes of each entity and relationship. The Relational Model is designed in accordance with the assumptions and constraints as mentioned in the document above.

Note: The arrows represent that a field is *derived* from another. For example, `productID` in `description` will contain values of `productID` from table `product`.


admin
adminID
hashed_password
customer
customerID
name
address
age
hashed_password
phone_number
supplier
supplierID
name
address
hashed_password
delivery_agent
daID
name
phone_number
availability
hashed_password
product
productID
supplierID
name
price
quantity
order
orderID
customerID
daID
status
ETA
product_reviews
reviewID
customerID
productID
stars
date
content
da_reviews
reviewID
customerID
daID
stars
date
content
description
productID
content
wallet
customerID
upiID
balance
consists_of
orderID
productID
quantity
delivered
daID
customerID
reviewID
purchased
productID
customerID
reviewID
cart
customerID
productID
quantity
sells
supplierID
productID
sold
supplierID
productID
date
quantity
Diagram Key
Primary Key
Foreign Key
Attribute
Arrow Key
customerID
supplierID
daID
productID
reviewID
orderID

Deadline 3

Implementing the Database Schema and Integrity Constraints, and Populating simulated data satisfying them

February 10, 2023

Iterations on the Database Schema

Tables added or renamed:

1. Since **address** is a composite attribute, we store it in a separate table since we would rarely need to search for the address. The search/join/union query efficiency would get affected if we store all the address data in the same stakeholder table.

2. The **consists_of** relationship table has been renamed to **order_product** to make it more intuitive.

3. Entity table **description** has been renamed to **product_description**, since `description` is a reserved keyword in MySQL. Its field **description** has also been renamed to **content** for the same reason.

4. Entity table **order** has also been renamed to **orders** for the same reason.

Columns/Fields added, removed, or renamed:

1. All fields called **hashed_password** have been renamed to **pwd** for simplicity.

2. Added field **email** to **customer**, **supplier**, and **delivery_agent** tables.

3. Added field **quantity** to table **product**.

4. Removed fields **reviewID** from the **product_review** and **da_review** tables, so as to decompose the ternary relationships **delivered** and **purchased**.

5. Added **order_date** and **delivery_date** to the **order** table.

Handling Different Attributes

Composite Attributes:

1. Addresses of all stakeholders are stored in a single table with a unique **addressID** assigned to each address. This **addressID** is then stored along with other stakeholder data in their tables.

2. Added sub-attribute **country** to the **address** composite attribute.

3. The **name** composite attribute is stored in the same table, since we might need to search for the names of the stakeholders. Each sub-attribute is kept as a column (**first_name**, **middle_initial**, **last_name**) in the table.

Multi-Valued Attributes:

1. Since **phone_number** is a multivalued attribute, we store it in a separate table, with the **phoneID** attribute from this table being stored in the **delivery_agent** and **customer** tables. This table does not have a primary key since each **phoneID** associates to one or more **num**'s (phone numbers).

Derived Attributes:

1. Changed **ETA** and **delivered** (earlier **status**) to derived attributes. **ETA** will be calculated as **order_date** + 15 days when required.

2. The attribute **delivered** will be `true` if **delivery_date** is not null, else `false`.

Since these constraints/relations will be implemented outside the database, and hence these fields are not present in the tables.

Assumptions:

1. The field **delivery_date** of **orders** is allowed to be `NULL`. This is because we do not know the **delivery_date** when the order is placed, and it is only updated when the order is delivered. We derive the attribute **delivered** from **delivery_date** to check if the order has been delivered or not.

2. Since **phone numbers** are implemented as multi-valued attributes, we assume that phone numbers are not unique to a customer. This means that multiple users may have the same phone number.

3. All primary keys are defined with the `AUTO_INCREMENT` constraints so that we do not need to insert ID values ourselves, and duplicity errors on primary keys are avoided.

4. **Product descriptions** and **review contents** are being stored using `TEXT` datatype in MySQL, which is not stored in the server memory and does not hamper query times.

5. All attributes that can not have null values have been specified as `NOT NULL`. Attributes like **last_name**, **middle_initial**, **content** in the **review** tables, and **delivery_date** in the **orders** table can have null values.

6. **Availability** of a delivery agent has been given a `DEFAULT` value of true, ie, when a delivery agent is added to the database, he is available to deliver an order by default.

7. **Balance** in a customer wallet, on account creation, has been given a `DEFAULT` value of 0.

8. **pwd**'s in all tables are stored as SHA1 hashes.
9. Some other integrity constrains have also been added, for example, **rating** must be between 0 and 5, and **quantity** and **price** must be positive.
10. Based on a rough idea of the types of queries we plan to use, some indices have also been added on the fields of the tables. These may be updated in the future.

Data Generation & Population

Data Generation

Most of the simulated data for the stakeholder and main entity tables was generated using https://www.mockaroo.com.

We utilised the (Ruby) code functionality to implement viable constraints on the data while data generation. The data for each table was downloaded as a CSV-file. We then used python scripts to generate the MySQL insertion queries and to populate relations. Data for some tables (like **cart** and **orders**) was mainly done through python scripts so as to make sure the existential constraints were not violated.

Data Population

All the tables of the database were pre-populated with data with integrity-constrains maintained to start querying. The database was populated with the following number of rows of data:

- `address`: 400 rows
- `phone_number`: 400 rows
- `admin`: 2 rows
- `supplier`: 200 rows
- `customer`: 200 rows
- `delivery_agent`: 200 rows
- `product`: 200 rows
- `orders`: 1000 rows
- `wallet`: 200 rows
- `product_review`: 200 rows
- `da_review`: 200 rows
- `cart`: 1073 rows
- `order_product`: 5552 rows

Deadline 4

SQL Queries and Relational Algebraic Operations

February 17, 2023

SQL Queries

We attempted to implement the most relevant queries for the application using the database. These queries are utilized by different stakeholders to perform their tasks. The following is a list of SQL Queries and their use cases:

1. The following queries are used while placing an order.

```
-- Find out an available delivery agent
SELECT daID FROM delivery_agent
WHERE availability = 1 ORDER BY daID ASC LIMIT 1;

-- Add the order to the table orders and assign the order to the
   selected delivery agent
INSERT INTO orders (customerID, daID, order_date)
VALUES (50, (SELECT daID FROM delivery_agent
WHERE availability = 1 ORDER BY daID ASC LIMIT 1), '2021-04-01');

-- Add the ordered products in the table order_product
INSERT INTO order_product (orderID, productID, quantity)
SELECT MAX(o.orderID), c.productID, c.quantity
FROM orders o INNER JOIN cart c ON o.customerID = c.customerID
WHERE o.customerID = 50
GROUP BY c.productID;

-- Reduce the quantity of products ordered from table product
UPDATE product p INNER JOIN cart c ON p.productID = c.productID
SET p.quantity = p.quantity - c.quantity
WHERE c.customerID = 50;

-- Delete the products from the cart
DELETE FROM cart WHERE customerID = 50;

-- Update the wallet balance of the customer
UPDATE wallet w
INNER JOIN orders o ON w.customerID = o.customerID
SET w.balance = w.balance - (
    SELECT SUM(p.price * op.quantity)
    FROM product p, order_product op
    WHERE p.productID = op.productID AND op.orderID = o.orderID
);

-- Update the availability of the delivery agent
UPDATE delivery_agent
SET availability = 0
WHERE daID = (
```

```
    SELECT daID FROM orders WHERE customerID = 50 AND orderID = (
        SELECT MAX(orderID) FROM orders WHERE customerID = 50
    )
);
```

2. This query is used to get a list of all suppliers who have all their products with average rating above 3.

```
SELECT s.supplierID, CONCAT(s.first_name, ' ', s.middle_initial, ' ',
    s.last_name) AS name
FROM supplier s
WHERE (
    SELECT AVG(pr.rating) FROM product_review pr, product p
    WHERE p.productID = pr.productID AND p.supplierID = s.supplierID
    GROUP BY p.supplierID
) > 3;
```

3. This query lists the suppliers who do not sell any products.

```
SELECT s.supplierID, CONCAT(s.first_name, ' ', s.middle_initial, ' ',
    s.last_name) AS name
FROM supplier s
WHERE NOT EXISTS (
    SELECT * FROM product p
    WHERE p.supplierID = s.supplierID
);
```

4. This query displays all the products that have average rating above or equal to 3.5 ordered by average rating.

```
SELECT p.productID, p.name, AVG(pr.rating) AS avg_rating
FROM product_review pr, product p
WHERE p.productID = pr.productID
GROUP BY p.productID
HAVING AVG(pr.rating) >= 3.5
ORDER BY AVG(pr.rating) DESC;
```

5. This query is used to display the top 40 delivery agents with the highest average rating.

```
SELECT da.daID, CONCAT(da.first_name, ' ', da.middle_initial, ' ',
    da.last_name) AS name, AVG(dr.rating) AS avg_rating
FROM da_review dr, delivery_agent da
WHERE da.daID = dr.daID
```

```
GROUP BY da.daID
ORDER BY AVG(dr.rating) DESC
LIMIT 40;
```

6. This query shows the undelivered orders of a delivery agent (where delivery date is NULL).

```
SELECT orderID, customerID, daID, order_date,
DATE_FORMAT(ADDDATE(order_date, INTERVAL 15 DAY), '%Y-%m-%d') AS ETA
FROM orders
WHERE daID = 1 AND delivery_date IS NULL;
```

7. This query is used to properly display the entire record of a customer.

```
SELECT
    customerID,
    CONCAT(first_name, ' ', middle_initial, ' ', last_name) AS name,
    CONCAT(apt_number, ', ', street_name, ', ', city, ', ', state, ' -
        ', zip, ', ', country) AS customer_address,
    age, ph.num as primary_phone_number, email
FROM
    customer, address,
    (
        SELECT num FROM phone_number, customer
        WHERE phone_number.phoneID = customer.phoneID
        AND customerID = 50 LIMIT 1
    ) AS ph
WHERE customer.addressID = address.addressID
AND customer.customerID = 50;
```

8. This query shows the top 15 customers who have spent the most money on their orders.

```
SELECT customer.customerID,
COUNT(orders.orderID) AS total_orders,
SUM(product.price * order_product.quantity) AS total_spent
FROM customer
INNER JOIN orders ON customer.customerID = orders.customerID
INNER JOIN order_product ON orders.orderID = order_product.orderID
INNER JOIN product ON order_product.productID = product.productID
GROUP BY customer.customerID
ORDER BY total_spent DESC
LIMIT 15;
```

9. This query is used to view the order history of a customer.

```
SELECT orders.orderID, orders.order_date, orders.delivery_date,
    product.name, order_product.quantity, product.price,
    (order_product.quantity * product.price) AS total_price,
    CASE
        WHEN (orders.delivery_date IS NULL) = True THEN 'Not Delivered'
        ELSE 'Delivered'
    END AS delivered
FROM orders
INNER JOIN order_product ON orders.orderID = order_product.orderID
INNER JOIN product ON order_product.productID = product.productID
WHERE orders.customerID = 1
ORDER BY orders.order_date;
```

10. This query is used to find out the total price of an order.

```
SELECT orderID, SUM(product.price * order_product.quantity) AS
    order_price
FROM order_product
INNER JOIN product ON order_product.productID = product.productID
WHERE order_product.orderID = 1;
```

11. This query is used to show all underlivered orders with their ETA for a customer.

```
SELECT orderID, order_date,
DATE_FORMAT(ADDDATE(order_date, INTERVAL 15 DAY), "%Y-%m-%d") AS ETA
FROM orders
INNER JOIN customer ON orders.customerID = customer.customerID
WHERE customer.customerID = 1 AND orders.delivery_date IS NULL
ORDER BY order_date;
```

12. This query find out the total revenue and total quantity sold per product for a supplier (sales statistics).

```
SELECT product.name, SUM(order_product.quantity) AS
    total_quantity_sold, SUM(order_product.quantity * product.price) AS
    total_revenue
FROM product
INNER JOIN order_product ON product.productID = order_product.productID
INNER JOIN orders ON order_product.orderID = orders.orderID
WHERE product.supplierID = 1
GROUP BY product.name;
```

13. This query is used to search through the product catalogue for names of the products using pattern matching.

```
SELECT name, AVG(rating)
FROM product
JOIN product_review ON product.productID = product_review.productID
WHERE name LIKE 'LED %'
GROUP BY name;
```

14. This query is used to add a product to a customer's cart.

```
INSERT INTO cart (customerID, productID, quantity) VALUES (50, 1, 1);
```

15. This query adds a new phone number into the table for a customer.

```
INSERT INTO phone_number VALUES
((SELECT phoneID FROM customer WHERE customerID = 50), '1234567890');
```

16. This query adds more quantity of products for an existing product.

```
UPDATE product SET quantity = quantity + 100 WHERE productID = 1;
```

17. This query updates the address of a customer.

```
UPDATE address SET
    apt_number = '100',
    street_name = 'Thomspon St.',
    city = 'Albany',
    state = 'New York',
    zip = '12207',
    country = 'United States'
WHERE addressID = (SELECT addressID FROM customer WHERE customerID = 1);
```

18. This query deletes a product from a customer's cart.

```
DELETE FROM cart WHERE customerID = 99 AND productID = 14;
```