BIG DATA FOR MANAGERS AND ANALYTICS

ENTITY-RELATION DIAGRAM FOR FLIPKART ECOMMERCE PLATFORM USING TOP-DOWN APPROACH

SUBMITTED TO:

PROF. ASHOK HARNAL

SUBMITTED BY:

MEGHA GARG (045030)
SHOBHIT GUPTA (045041)
SANSKRITI BAHL (045050)

INTRODUCTION

This report presents a comprehensive analysis of the Entity-Relationship Diagram (ERD) created for Flipkart, a leading e-commerce platform in India. The ERD was developed using a top-down approach to model the core entities and relationships within Flipkart's database system.

E-commerce platforms like Flipkart require robust and scalable database designs to handle vast amounts of data, from user information and product catalogs to order processing and inventory management. The ERD serves as a crucial blueprint for designing and implementing such a database system, ensuring efficient data storage, retrieval, and manipulation.

In this project, we focused on creating a streamlined yet comprehensive ERD that captures the essential components of Flipkart's operations. This design aims to support key functionalities such as user management, product categorization, order processing, payment handling and order details

In developing an optimized database system for an e-commerce platform like Flipkart, a top-down approach is employed to ensure a robust and scalable architecture. The process begins by identifying high-level requirements, focusing on the core entities such as customers, products, orders, and payments, along with their key relationships and functionalities necessary to support seamless operations. This initial phase is followed by the conceptual design, where an Entity-Relationship (ER) diagram is created to map out these entities and their interactions, abstracting away the complexities of data storage. The model is then normalized, decomposing it into smaller entities to eliminate redundancy and avoid data anomalies while maintaining logical consistency. Next, the conceptual model is translated into a logical design, where it is represented as a set of normalized tables and relationships, independent of any specific database management system (DBMS). In the final phase, the logical design is transformed into a physical design tailored to the chosen DBMS, involving the definition of indexes, storage parameters, and other implementation-specific details. This structured approach ensures that the database system is not only aligned with Flipkart's operational needs but is also flexible enough to support future growth and technological advancements.

OBJECTIVES

The primary objectives of this project were to:

- Design a comprehensive ERD representing Flipkart's e-commerce ecosystem
- Identify key entities and their relationships using top down approach in database system.
- Establish a foundation for database implementation for Flipkart
- Provide insights into the platform's data structure

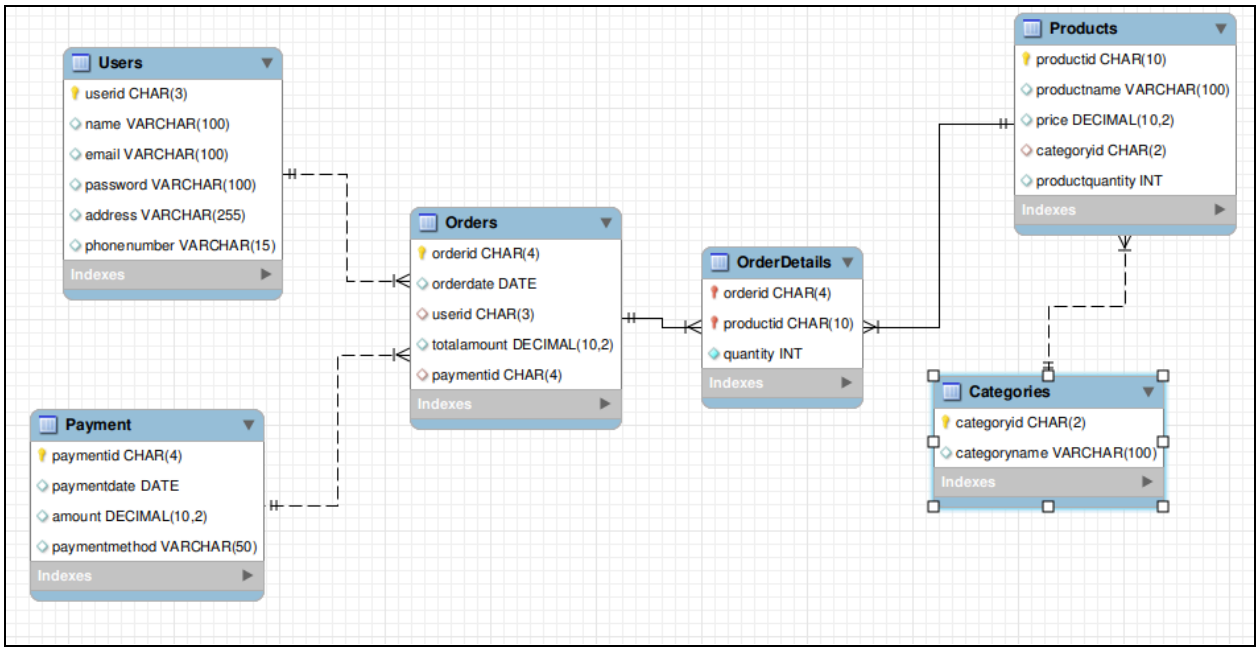
IDENTIFYING HIGH-LEVEL REQUIREMENTS

A comprehensive analysis was undertaken to identify the high-level requirements of Flipkart's e-commerce platform involving understanding the core functionalities that the database system must support, including **user management, product cataloging, order processing, payment handling, and order details** The focus is on recognizing the primary entities—such as Users, Products, Orders, Payments, Categories, Order Detail—and their interrelationships to ensure that the database design aligns with the operational needs and business objectives of the platform.

Determined attributes for each entity: For each entity, we defined relevant attributes. For example, Users have attributes like **user ID, name, email, password, address, and phone number.**

Established relationships between entities and analyzing how these entities interact within the system. For instance, we established that Users place Orders, Products belong to Categories, and Orders are linked to Payments.

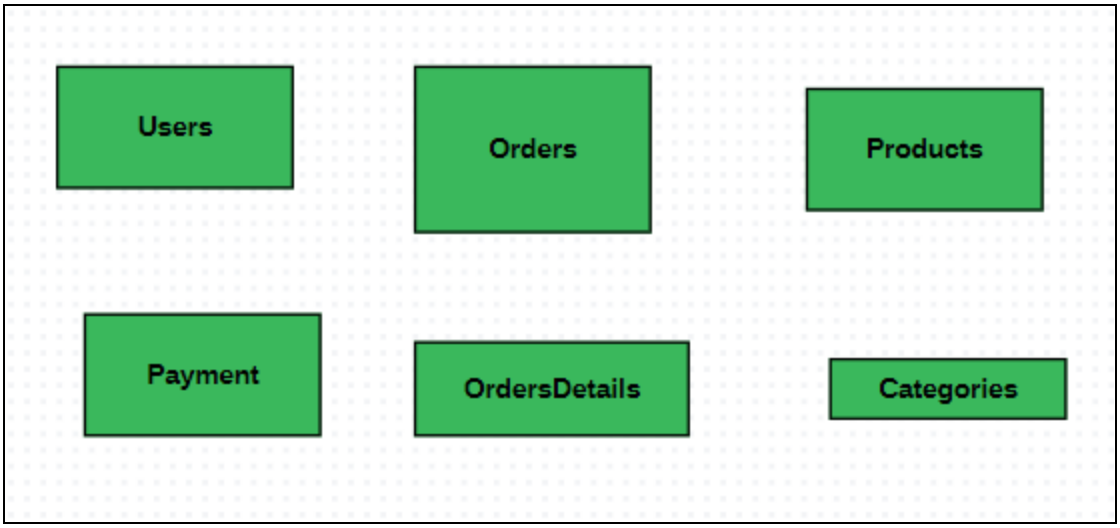
ENTITY RELATIONSHIP DIAGRAM ON MYSQL WORKBENCH



PROCESS

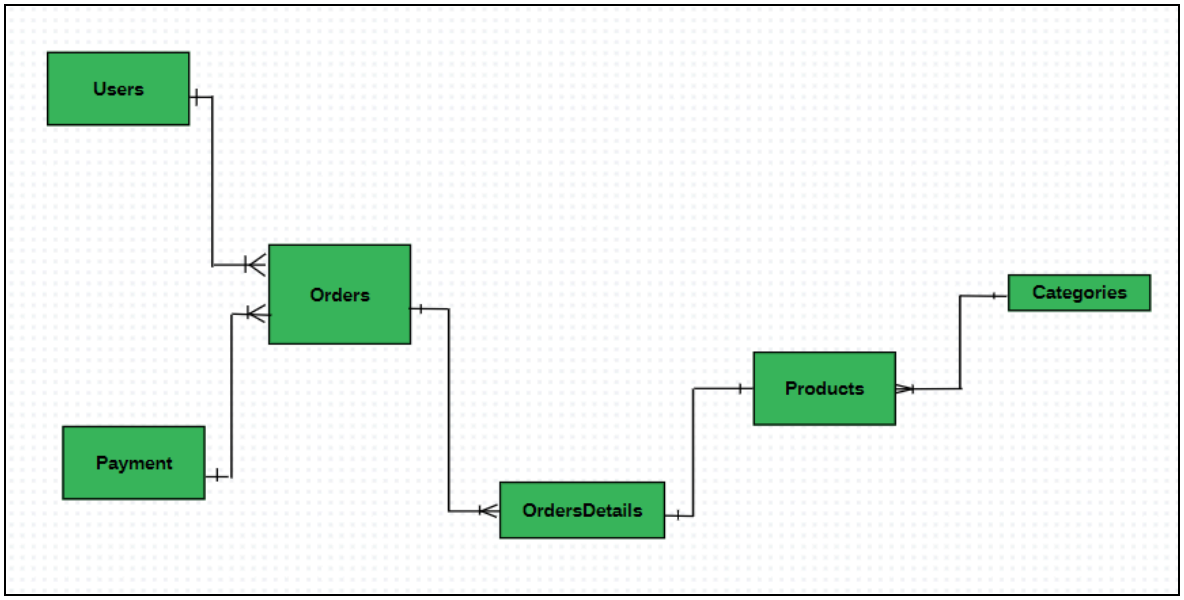
Identify the major entities

In the top-down design of an ERD for a Flipkart-like e-commerce platform, the first step is to identify the major entities essential to the system. These entities include Users, Orders, Products, Categories, Payment, and OrderDetails. Each entity represents a fundamental component of the platform: Users correspond to the individuals interacting with the platform, Orders capture the transactions made by these users, Products represent the items available for purchase, Categories organize these products into distinct groups, Payment details the financial transactions, and OrderDetails provide a breakdown of the specific products within each order.



Defining Relationship

The relationships between these entities are carefully established to model the real-world operations of an e-commerce platform. For example, a single user can place multiple orders, forming a one-to-many relationship between Users and Orders. Each order can contain multiple products, which creates a one-to-many relationship between Orders and OrderDetails. Products belong to specific categories, establishing a many-to-one relationship between Products and Categories. Each order is associated with a single payment record, forming a one-to-one relationship between Orders and Payment. These relationships are integral to the database's ability to represent and manage the platform's operations accurately.

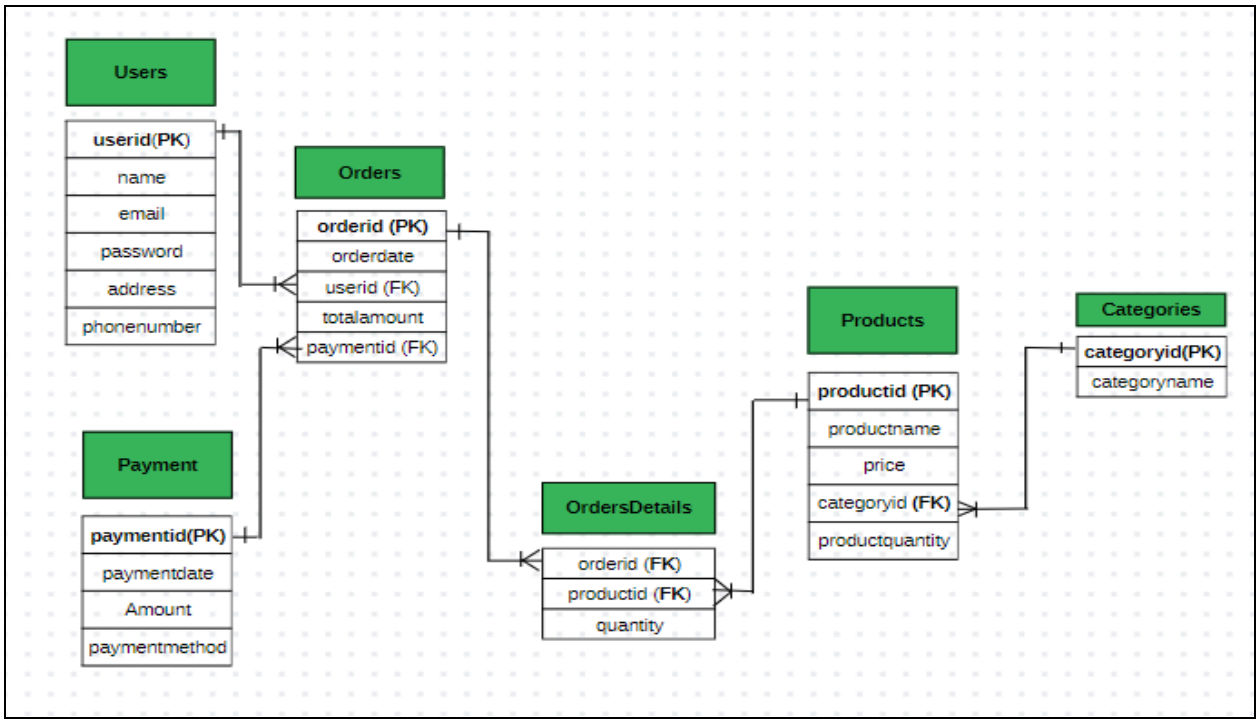


Defining attributes

After identifying the entities, **attributes are defined** for each to capture necessary details without redundancy. For instance, the Users entity includes attributes such as `userid`, `name`, `email`, `password`, `address`, and `phonenumber`, ensuring a comprehensive profile for each user etc.

Identifying primary key and foreign key

In relational database design, primary keys and foreign keys are used to ensure data integrity and maintain relationships between tables. Each entity has a primary key, such as `userid` in Users, `orderid` in Orders, and `productid` in Products, uniquely identifying each record. Foreign keys establish links between entities, like `userid` in Orders referencing the `userid` in Users, or `categoryid` in Products referencing `categoryid` in Categories. These relationships ensure consistent and valid interactions between the data.



SQL IMPLEMENTATION

- **Translation of ERD into SQL Code**: The visual representation of the ERD was meticulously converted into SQL statements, ensuring an accurate translation of the conceptual design into a tangible database structure. This step laid the groundwork for the database by defining the entities and their relationships in SQL.
- **Creation of Tables for Each Entity**: SQL CREATE TABLE statements were crafted for each identified entity, including Users, Products, Categories, Orders, Payments, and Order Detail. These statements were designed to establish the necessary tables with their respective attributes, as outlined in the ERD.
- **Definition of Primary and Foreign Keys**: Primary keys were specified for each table to uniquely identify records within the database. Additionally, foreign keys were defined where relationships existed between tables, linking related data and establishing the foundational structure of the database.

- **Establishment of Relationships Through Foreign Key Constraints**: Foreign key constraints were implemented to maintain referential integrity between related tables. This step was crucial in ensuring that relationships, such as those between Users and Orders or Products and Categories, were properly enforced within the database.
- **Verification of Relationships in MySQL Workbench**: Each relationship was confirmed using MySQL Workbench to ensure that the foreign key constraints were correctly established and functioning as intended. This verification process was essential in validating the integrity and accuracy of the database design.

DATA ACCESS LANGUAGE

Data Access Language (DAL) is a set of commands or languages used to interact with databases, enabling users to retrieve, insert, update, and delete data. It serves as the interface for managing and accessing the data stored within a database system, ensuring that information can be efficiently manipulated and retrieved as needed.

1. **User Management**

- **Entities Accessed**: Users
- **Data Accessibility:** Full access to user data including name, email, password (hashed/encrypted), address, and phone number.
- Managing user profiles, authentication, and customer support.
- **User:** *user_mgmt*
- **Password:** *user_mgmt_password*
- **Code:**

CREATE USER 'user_mgmt'@'localhost' IDENTIFIED BY 'user_mgmt_password';

GRANT SELECT, INSERT, UPDATE, DELETE ON flipkart.Users TO 'user_mgmt'@'localhost';

FLUSH PRIVILEGES;

2. **Product Management User**

- Entities Accessed: Categories and Products
- Data Accessibility: Full access to categories and product data including product names, prices, quantities, and associated categories.
- Managing product listings, categorization and inventory management.
- **User:** *product_mgmt*
- **Password:** *product_mgmt_password*
- **Code:**

CREATE USER 'product_mgmt'@'localhost' IDENTIFIED BY 'product_mgmt_password';
GRANT SELECT, INSERT, UPDATE, DELETE ON flipkart.Categories TO 'product_mgmt'@'localhost';

GRANT SELECT, INSERT, UPDATE, DELETE ON flipkart.Products TO 'product_mgmt'@'localhost';

FLUSH PRIVILEGES;

3. **Sales and Order Management User**

- Entities Accessed: Orders and OrderDetails
- Data Accessibility: Full access to order details including order ID, order date, user ID, total amount, and associated products in each order.
- Limited access to user data: Only user IDs to track orders.
- **User:** *sales_mgmt*
- **Password:** *sales_mgmt_password*
- **Code:**

```
CREATE USER 'sales_mgmt'@'localhost' IDENTIFIED BY 'sales_mgmt_password';

GRANT SELECT, INSERT, UPDATE, DELETE ON flipkart.Orders TO 'sales_mgmt'@'localhost';

GRANT SELECT, INSERT, UPDATE, DELETE ON flipkart.OrderDetails TO 'sales_mgmt'@'localhost';

GRANT SELECT ON flipkart.Users(userid) TO 'sales_mgmt'@'localhost';

FLUSH PRIVILEGES;
```

4. **Logistics and Inventory User**

- Entities Accessed: Products and OrderDetails
- Data Accessibility: Access to product data (quantities and IDs) to manage inventory and coordinate shipments. Access to order details to ensure correct items are shipped and delivered.
- Managing inventory levels, ensuring timely shipments, and handling returns.
- **User:** *logistics_mgmt*
- **Password:** *logistics_mgmt_password*

- **Code:**

CREATE USER 'logistics_mgmt'@'localhost' IDENTIFIED BY 'logistics_mgmt_password';

GRANT SELECT, UPDATE ON flipkart.Products TO 'logistics_mgmt'@'localhost';

GRANT SELECT ON flipkart.OrderDetails TO 'logistics_mgmt'@'localhost';

FLUSH PRIVILEGES;

5. **Analytics and Reporting User**

- Entities Accessed: All Entities (Aggregated Data)
- Data Accessibility: Aggregated and anonymized data from all tables to generate insights, trends, and performance metrics.
- Analyzing sales, user behavior, product performance, and generating reports for decision-making.
- **User:** *analyst*
- **Password:** *analyst_password*
- **Code:**

CREATE USER 'analyst'@'localhost' IDENTIFIED BY 'analyst_password';

GRANT SELECT ON flipkart. TO 'analyst'@'localhost';*

FLUSH PRIVILEGES;

INSIGHTS

1.1 Core Entities

The ERD reveals six core entities: Users, Categories, Products, Payment, Orders, and Order Detail. Each entity plays a crucial role in the e-commerce ecosystem:

- Users: Represents customer information
- Categories: Allows for logical grouping of products
- Products: Contains details of items available for purchase
- Payment: Tracks transaction details
- Orders: Records customer purchases
- Order Details: Junction Table for Product and Order

1.2 Relationships

- Users can place multiple Orders (one-to-many): This allows a single user to have a history of multiple orders over time.
- Products belong to Categories (many-to-one): Enables efficient categorization and browsing of products.
- Orders are associated with one Payment (one-to-one): Ensures each order has a corresponding payment record.
- Products are tracked in Categories (one-to-many): Facilitates real-time updates for highest selling categories in turn use for stock management.

1.3 Data Integrity

Foreign key constraints ensure referential integrity between related tables. For example, an order cannot exist without a corresponding user, maintaining data consistency.

1.4 Scalability

The design allows for easy addition of new products, categories, and users. This is crucial for an e-commerce platform that continuously expands its user base and product offerings.

1.5 Payment Tracking

The separate Payment table enables detailed transaction recording and supports various payment methods. This design choice allows for easy integration of new payment options in the future.

1.6 Inventory Management

The Category table and quantity attribute facilitates stock tracking and can be used for reordering processes. This is essential for maintaining accurate product availability and managing supply chain operations.

CONCLUSION

The Entity-Relationship Diagram created for Flipkart's e-commerce platform provides a solid foundation for a robust and scalable database system. By capturing the essential entities (Users, Products, Categories, Orders, Payments, and Order Detail) and their interrelationships, the ERD effectively models the core functionalities of an e-commerce ecosystem.

The design demonstrates several strengths:

1. Flexibility: It can accommodate growth in users, products, and transactions.

2. Data Integrity: Relationships and constraints ensure consistent and reliable data.

3. Operational Support: The structure facilitates key e-commerce operations from user management to inventory control.

The SQL implementation of this ERD provides a practical starting point for database creation. However, as Flipkart continues to evolve, this model may need to be expanded. Future enhancements could include:

- Integration of customer reviews and ratings
- Support for wish lists and saved items
- Incorporation of seller information for a marketplace model
- Advanced inventory management features

In conclusion, this ERD serves as a valuable tool for understanding and implementing Flipkart's data architecture. It provides a clear visualization of the system's structure, guiding database design and application development. As Flipkart grows and adapts to changing market demands, this foundational design can be iteratively refined to support new features and optimizations, ensuring the platform's continued success in the competitive e-commerce landscape.