# Database Management For Analytics

# E-commerce Database Management System

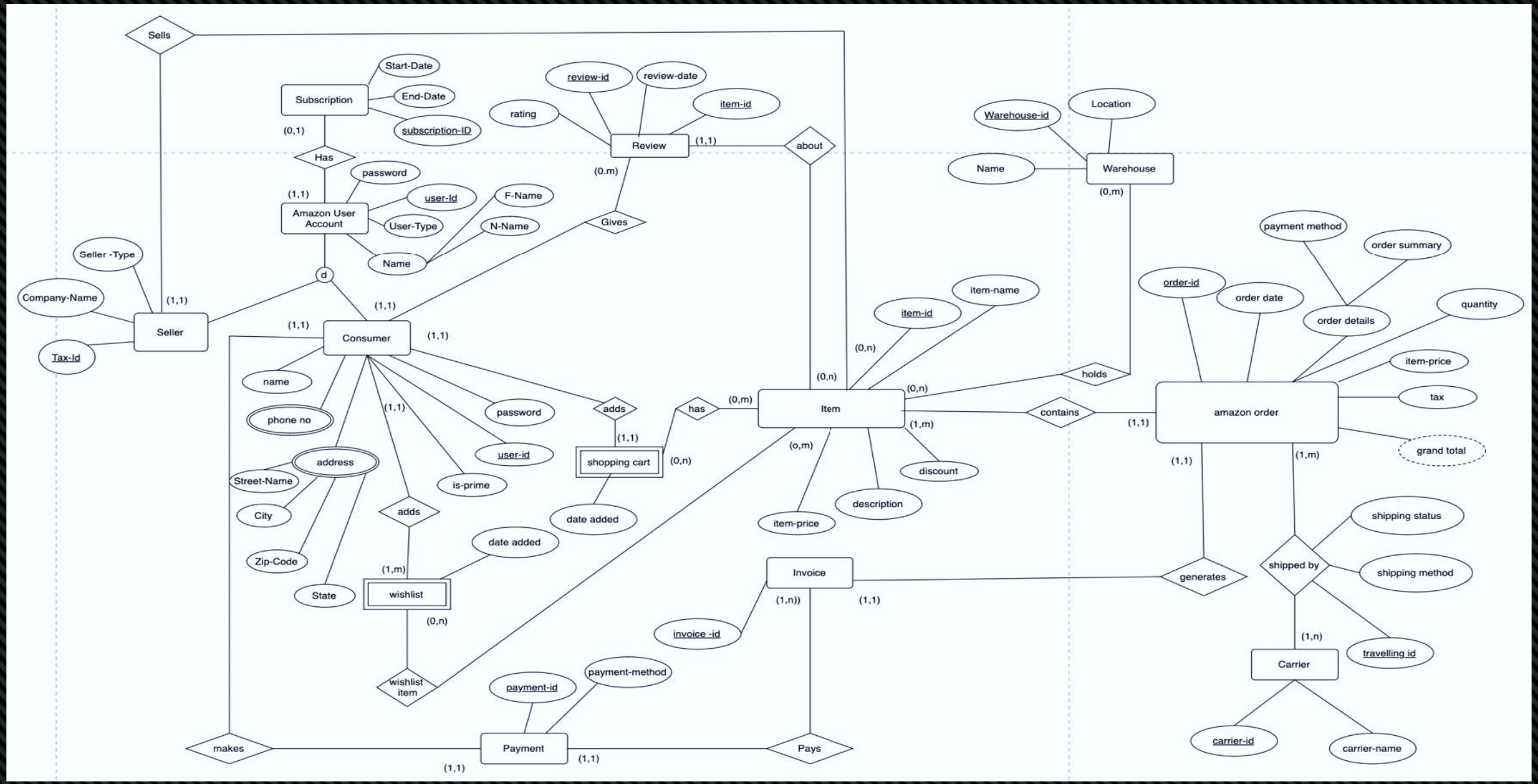

**PROJECT BY :**

**GROUP 15**

# PROBLEM STATEMENT

Our objective is to analyze an Amazon dataset, identifying customer segments, behavioral trends, and forecasting customer attrition. This analysis aims to tailor services for personalized customer experiences, answering key questions about visitation patterns, expenditure, and preferred items.

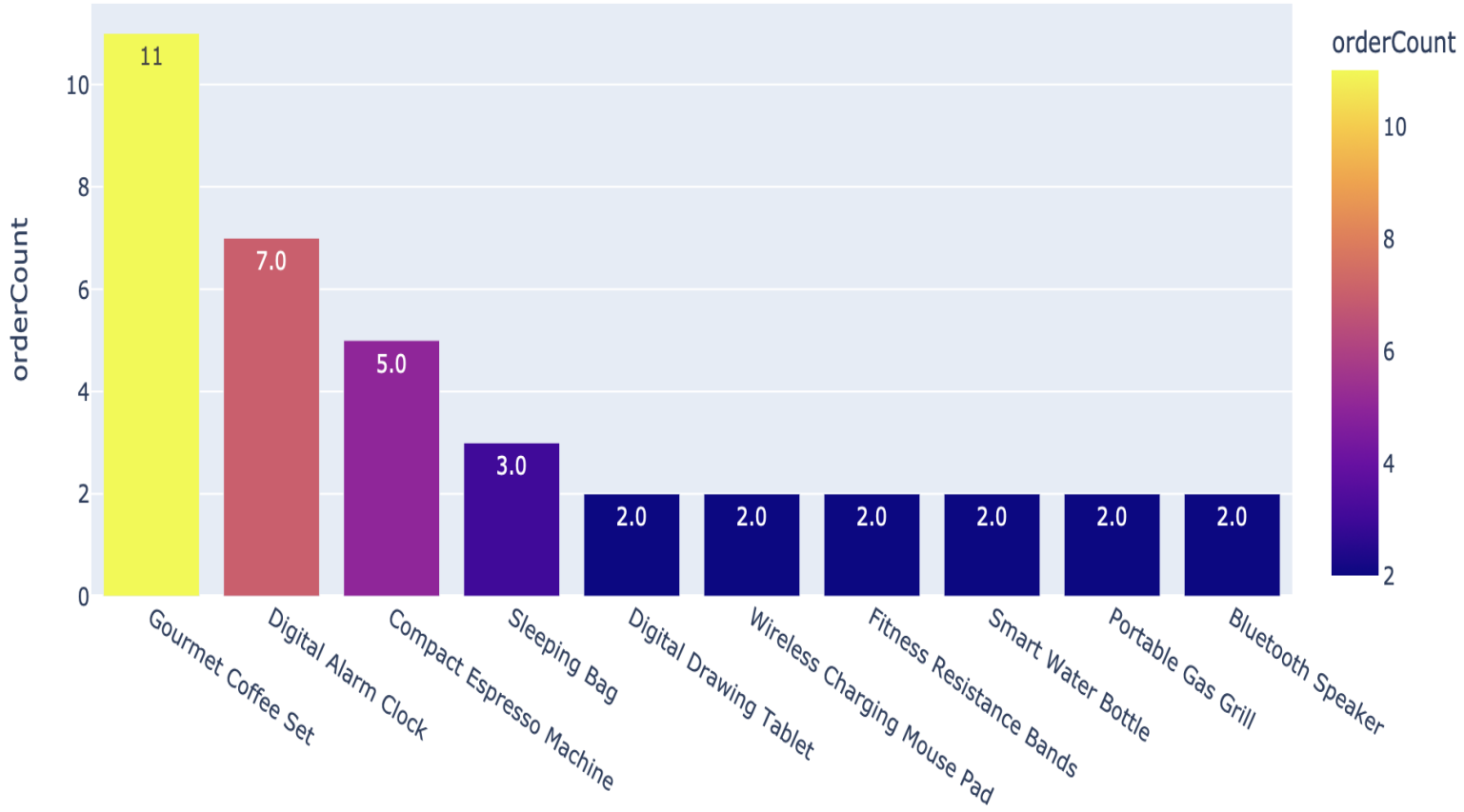
# Enhanced Entity Relationship Diagram for an Ecommerce Platform



# SCOPE OF ANALYTICS

1. **Visitation Patterns:**
    - Analyze consumer visitation data over time.
    - Identify peak visitation periods.
2. **Expenditure Analysis:**
    - Evaluate consumer spending patterns.
    - Identify high-value customers.
3. **Preferred Items:**
    - Determine popular items through purchase history.
    - Tailor recommendations based on preferences.

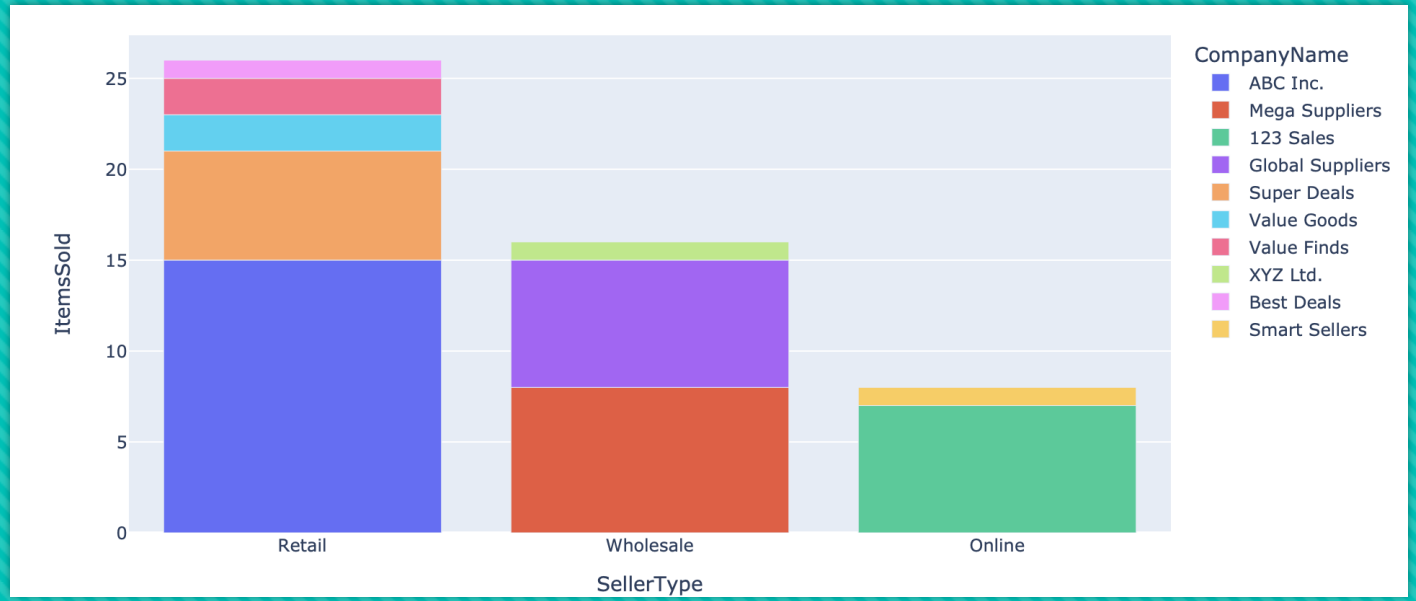
```
sql_query = """
SELECT i.item_name, COUNT(io.Itemid) AS orderCount
FROM item AS i
left JOIN item_order AS io ON i.itemid = io.itemid
GROUP BY i.item_name
ORDER BY orderCount DESC
LIMIT 10;
"""
query1 = pd.read_sql_query(sql_query, connection)
```




**Top 10 Frequently Ordered Items:**

- Highlight essential everyday products like coffee sets and alarm clocks.
- Emphasize the consistent popularity and demand for these items.

```
sql_query = """
SELECT
  s.UserID AS SellerID,
  s.CompanyName,
  s.SellerType,
  COUNT(i.ItemID) AS ItemsSold,
  SUM(i.quantity) AS TotalQuantitySold
FROM
  seller s
LEFT JOIN
  item i ON s.UserID = i.UserID
WHERE
  i.Order_ID IS NOT NULL
GROUP BY
  s.UserID, s.CompanyName, s.SellerType
ORDER BY
  ItemsSold DESC
  limit 10;
"""
query2 = pd.read_sql_query(sql_query, connection)
```




**1.Strategic Decision Support:**

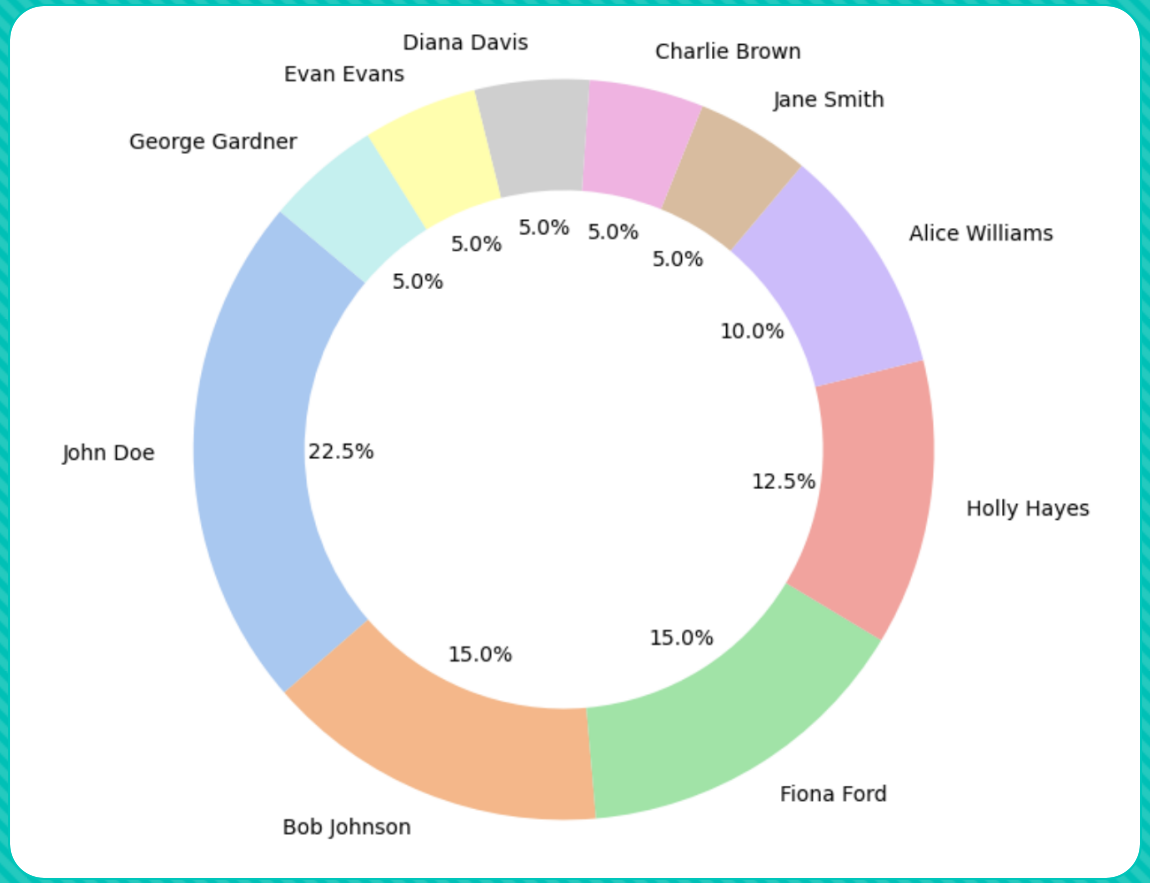
1. SQL query offers a comprehensive view of seller data, meticulously tracking total items sold.
2. Differentiation between wholesale, retail, and online sellers provides insights for strategic decision-making

**2.Transparent Performance Evaluation:**

1. Data-driven approach underscores platform commitment to transparency and performance assessment.
2. Nuanced understanding of relationships between seller types reveals that Retail sellers lead in both variety and quantity of items sold, shaping the overall transactional landscape.

```
sql_query = """
SELECT
  c.userID AS ConsumerID,
  c.cons_name AS ConsumerName,
  COUNT(o.Order_ID) AS TotalOrders,
  MIN(o.order_date) AS FirstOrderDate,
  MAX(o.order_date) AS LatestOrderDate,
  DATEDIFF(MAX(o.order_date), MIN(o.order_date)) AS DaysBetweenFirstAndLatestOrder
FROM
  consumer c
JOIN
  Amazon_Order o ON c.userID = o.userID
GROUP BY
  c.userID, c.cons_name
ORDER BY
  TotalOrders DESC
  limit 10;

"""
query4 = pd.read_sql_query(sql_query, connection)
```




**1.Top 10 High Valued Customers :**

- Highlight the significant impact of top customers on the digital marketplace.
- Showcase their penchant for the latest trends, contributing to the vibrancy of the platform.

**2.Nuanced User Spending Patterns:**

- Correlated queries delve into average grand total spent by each user.
- Uncover distinctive spending habits, emphasizing the diverse preferences and unique contributions of top customers to the financial dynamics of the online marketplace.

# LIVE DEMO OF SQL and NoSQL IMPLEMENTATION