Data Analyst Internship - Task 3: SQL for Data Analysis

Suchismita Maity

November 17, 2025

Abstract

This report fulfills the requirements for Task 3: SQL for Data Analysis. It includes a series of SQL queries demonstrating data extraction and analysis on a sample e-commerce database. The queries cover all topics mentioned in the mini-guide, including SELECT statements, JOINs, subqueries, aggregate functions, views, and indexes.

1 SQL Queries and Simulated Outputs

This section details each query as required by the task hints.

1.1 a. SELECT, WHERE, ORDER BY, GROUP BY

1.1.1 SELECT & WHERE

Query: Find all customers from 'New York'.

```
SELECT CustomerName, Email, City FROM Customers WHERE City = 'New York';
```

Simulated Output:

CustomerName	Email	City
Alice Smith	alice@email.com	New York
Charlie Brown	charlie@email.com	New York

1.1.2 ORDER BY

Query: List all products, ordered by price from high to low.

```
SELECT ProductName, Category, Price FROM Products ORDER BY Price DESC;
```

Simulated Output:

ProductName	Category	Price
Laptop	Electronics	1200.00
Smartphone	Electronics	700.00
Desk Chair	Furniture	150.00
Coffee Maker	Appliances	80.00

1.1.3 GROUP BY

Query: Count the number of customers in each city.

```
SELECT City, COUNT(CustomerID) AS NumberOfCustomers FROM Customers GROUP BY City;
```

Simulated Output:

City	NumberOfCustomers
New York	2
Chicago	2
San Francisco	1

1.2 b. Use JOINS (INNER, LEFT, RIGHT)

1.2.1 INNER JOIN

Query: Get a list of customers and their order dates.

```
SELECT C.CustomerName, O.OrderID, O.OrderDate FROM Customers C INNER JOIN Orders O ON
C.CustomerID = O.CustomerID;
```

Simulated Output:

CustomerName	OrderID	OrderDate
Alice Smith	1001	2025-11-10
Bob Johnson	1002	2025-11-11
Alice Smith	1003	2025-11-12
Charlie Brown	1004	2025-11-12

1.2.2 LEFT JOIN

Query: List all customers and any orders they have placed (includes customers with no orders).

```
SELECT C.CustomerName, O.OrderID FROM Customers C LEFT JOIN Orders O ON C.CustomerID =
O.CustomerID;
```

Simulated Output:

CustomerName	OrderID
Alice Smith	1001
Alice Smith	1003
Bob Johnson	1002
Charlie Brown	1004
David Lee	*NULL*
Eve Williams	*NULL*

1.3 c. Write subqueries

1.3.1 Subquery in WHERE

Query: Find customers who have placed an order.

```
SELECT CustomerName, Email FROM Customers WHERE CustomerID IN (SELECT DISTINCT CustomerID
FROM Orders);
```

Simulated Output:

CustomerName	Email
Alice Smith	alice@email.com
Bob Johnson	bob@email.com
Charlie Brown	charlie@email.com

1.4 d. Use aggregate functions (SUM, AVG)

1.4.1 SUM

Query: Calculate the total revenue from all orders.

```
SELECT SUM(TotalAmount) AS TotalRevenue FROM Orders;
```

Simulated Output:

TotalRevenue
2210.00

1.4.2 AVG

Query: Calculate the average price of products in the 'Electronics' category.

```
SELECT AVG(Price) AS AverageElectronicsPrice FROM Products WHERE Category = 'Electronics';
```

Simulated Output:

AverageElectronicsPrice
950.00

1.5 e. Create views for analysis

1.5.1 CREATE VIEW

Query: Create a view that joins customer and order information.

```
CREATE VIEW CustomerOrderSummary AS SELECT C.CustomerName, C.Email, C.City, O.OrderID,
O.OrderDate, O.TotalAmount FROM Customers C JOIN Orders O ON C.CustomerID = O.CustomerID;
```

Simulated Output:

```
View 'CustomerOrderSummary' created successfully.
```

1.5.2 Query the View

Query: Select from the new view.

```
SELECT * FROM CustomerOrderSummary WHERE City = 'New York';
```

Simulated Output:

CustomerName	Email	City	OrderID	OrderDate	TotalAmount
Alice Smith	alice@...	New York	1001	2025-11-10	1280.00
Alice Smith	alice@...	New York	1003	2025-11-12	150.00
Charlie Brown	charlie@...	New York	1004	2025-11-12	80.00

1.6 f. Optimize queries with indexes

1.6.1 CREATE INDEX

Query: Create an index on the CustomerName column.

```
CREATE INDEX idx_customer_name ON Customers(CustomerName);
```

Simulated Output:

```
Index 'idx_customer_name' created successfully.
```

1.6.2 Query using Index

Query: This query would now be faster.

```
SELECT * FROM Customers WHERE CustomerName = 'Alice Smith';
```

Simulated Output:

CustomerID	CustomerName	Email	City
1	Alice Smith	alice@email.com	New York