**Explanation:**

**Exercise 1: Writing Queries**

**2. Optimize the above query to ensure it is efficient, considering indexes and other**

**optimization techniques.**

a) check for the columns that are used in both join and filter are properly indexed

b) only select required columns

c) Use of With clause/common table expression for filtering out data of order table and doing aggregation by customerid to reduce the number of records to join with customer table.

i. CREATE INDEX idx\_orders\_customer\_date ON Orders(customer\_id);

ii. CREATE INDEX idx\_orders\_date ON Orders(order\_date)

ii. CREATE INDEX idx\_order\_details\_order\_id ON Order\_Details(order\_id);

**2.Exercise 2: SQL Query Optimization**

Generally when we have to do optimization we first check for indexes,IN and required column selection. When we do all these changes then we check for explain plan to check for a full table scan.

a) Index creation on all the join columns and filtered columns.

CREATE INDEX idx\_order\_details\_orderpro\_id ON Order\_Details(product\_id);

**3.Exercise 3: Data Modeling**

a) Star schema : Star schema used primarily in data warehousing. It involves a central fact table connected to one or more dimension tables.

b) Snowflake schema : Snowflake Schema is a more normalized version of the star schema. It involves normalized dimension tables that are connected to the central fact table.

c) 3NF : 3NF is a relational database design approach that aims to reduce redundancy and improve data integrity by organizing data into tables that are in the third normal form.

The inventory management system focuses on operational transactions and data integrity so I choose 3NF because it ensures that the database is well organized and minimizes redundancy and ensures data integrity.This is crucial for transactional systems where the accuracy of sales, product, and supplier information is vital.

3NF is typically used in OLTP systems for operational efficiency.

Star and Snowflake schema mainly used for Analytical and reporting purposes. They are suitable for OLAP systems.

**4. Exercise 4: Creating DBT Models**

To ensure the quality of data in these model we create schema level validation check as well as we create custom tests.

Schema.yml → this includes basis checks like unique,notnull,email format validation,phone number validation.

Custom.yml → in which we define a custom test scenario and add the name in the schema.yml file.

**5. Data Analysis and Dashboard in PowerBI**

I don’t have experience in power bi but I still tried to cover the scenarios.

I have used pie charts for sales by category and sales by region because I want to show the percentage, not the exact number. When I have to show the exact number we can go for a bar chart.

For month sales comparison I have used line charts to show how sales are moving from month to month.