Expectations: The candidate is expected to attempt the below assessment. Outcome is an sql query to achieve the mentioned objectives.

Schema

Table 1: Users

Userid *(primaryKey)*	INT
Company Name	Varchar
City	Varchar
State	Varchar
Signup Date	BIGINT

Table 2: Orders

Orderid *(primaryKey)*	INT
Order Date	BIGINT
Pairs	INT
Order Value	Double(10,2)
Userid *(foreignKey)*	INT
sellerID (foreign Key)	INT
TransactionId *(foreignKey)*	INT
Order Ready Date	BIGINT
Payment Type	Varchar (COD, Prepaid)
Delivery Date	BIGINT
Status	Varchar (Placed, Confirmed, Dispatched, Intransit, Delivered, RTO, Cancelled, Lost)

Table 3: order_details (SKU level details within an order)

OrderDetailID	INT
Orderid	INT
Order Date	BIGINT
SKUID	INT
Pairs	INT
Orderdetail_Value	Decimal

Table 4: Returns

DisputeId *(primaryKey)*	INT
ReturnBookedOn	BIGINT
ReturnType	INT
orderId *(foreignKey)*	INT
RefundValue	INT
BuyerRefundedOn	BIGINT

*** ReturnType=1 is Genuine, ReturnType=2 is NonGenuine*

Section 1: Basic analytics

Q1. SKU Level sales month on month

Q2. Most sold SKUs in maharashtra

Q3. State with highest % of orders as COD orders

Q4. Find the percentage of returns booked out of the total orders placed in the last 30 Days?
Exp Result: Percentage of Returns

Q5. All churned buyers (Buyers who have not placed a order in the last 30 days)
Exp Result: Buyer ID, Last order date, Last delivery Date

Q6. Find the total Users who have received refunds of more than Rs.10000 in the last 30 days?
Exp result - userid,Total Orders Placed (lifetime), Total Returns Booked, Last Refund Date, Total Refund Value, Total No of Genuine Returns, Total No of Non-Genuine Returns

Q7. SKU with most returns bifurcated into Genuine and Non Genuine
Exp Result: SKUID, % of orders returned, Genuine contribution, Non Genuine contribution

Q8. Buyer level Recency, Frequency analysis: To be able to figure out the avg frequency at which a buyer purchases and when was the last

Section 2: Advanced Data analytics

Q.1 Bucketing users into Large, Medium and small. Size is derived based on the sales contribution of the user at a platform level. Definitions: (In terms of order Value)

- Large: TOP 25% contributors
- Medium: Next 50%
- Small: Last 25%

Expected Outcome: UserID, % Contribution to the platform sales and Size-bucket

Q2. Cohort analysis: A Cohort can be defined as a set of users who place their 1st order in the same month and analyse their count month on month. Expected outcome:

1st order Month	M0	M1	M2	M3	M4
Jan'20	Count of users who placed there 1st order in Jan	Count of users from M0 who placed an order in Feb	Count of users from M0 who placed an order in Mar	Count of users from M0 who placed an order in Apr	Count of users from M0 who placed an order in May
Feb'20	Count of users who placed there 1st order in Feb	Count of users from M0 who placed an order in Mar	Count of users from M0 who placed an order in Apr	Count of users from M0 who placed an order in May	Count of users from M0 who placed an order in Jun
Mar'20	Count of users who placed there 1st order in Mar	Count of users from M0 who placed an order in Apr	Count of users from M0 who placed an order in May	Count of users from M0 who placed an order in Jun	Count of users from M0 who placed an order in Jul

Q3. A seller is considered to have breached SLA if the order placed and order ready time stamps are greater than 3 days. Find out the top 10 percentile of sellers causing delays on a platform level. Expected Outcome: Seller ID, contribution on platform, Delay % of seller