Task 3: SQL for Data Analysis

1) Use SELECT, WHERE, ORDER BY, and GROUP BY

```
SELECT `Gender`, AVG(`Purchase Amount (USD)`) AS `AvgSpent`
FROM shopping_trends_updated
WHERE `Season` = 'Spring'
GROUP BY `Gender`
ORDER BY `AvgSpent` DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Gender	AvgSpent
Female	58.7911
Male	58.7130

2) Use JOINS (INNER, LEFT, RIGHT)

```
-- LEFT JOIN
SELECT s.`Customer ID`, s.`Shipping Type`, c.`Cost`
FROM shopping_trends_updated s
LEFT JOIN shipping_cost c
ON s.`Shipping Type` = c.`Shipping Type`;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Customer ID	Shipping Type	Cost
1	Express	10
2	Express	10
3	Free Shipping	0
4	Next Day Air	20
5	Free Shipping	0
6	Standard	NULL
7	Free Shipping	0
8	Free Shipping	0
9	Express	10

Result 3

Task 3: SQL for Data Analysis

3) Write subqueries

```
25 •  SELECT *
26    FROM shopping_trends_updated
27    WHERE `Purchase Amount (USD)` > (
28        SELECT AVG(`Purchase Amount (USD)`)
29        FROM shopping_trends_updated
30    );
31
32
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Customer ID	Age	Gender	Item Purchased	Category	Purchase Amount (USD)	Location	Size	Color	Season	Review Rating	Subscription Status	Shipping Type
2	19	Male	Sweater	Clothing	64	Maine	L	Maroon	Winter	3.1	Yes	Express
3	50	Male	Jeans	Clothing	73	Massachusetts	S	Maroon	Spring	3.1	Yes	Free Shipping
4	21	Male	Sandals	Footwear	90	Rhode Island	M	Maroon	Spring	3.5	Yes	Next Day Air
7	63	Male	Shirt	Clothing	85	Montana	M	Gray	Fall	3.2	Yes	Free Shipping
9	26	Male	Coat	Outerwear	97	West Virginia	L	Silver	Summer	2.6	Yes	Express
12	30	Male	Shorts	Clothing	68	Hawaii	S	Olive	Winter	4.9	Yes	Store Pickup
13	61	Male	Coat	Outerwear	72	Delaware	M	Gold	Winter	4.5	Yes	Express
16	64	Male	Skirt	Clothing	81	Rhode Island	M	Teal	Winter	2.8	Yes	Store Pickup

shopping_trends_updated 4

4) Use aggregate functions (SUM, AVG)

```
32 •  SELECT `Location`, COUNT(*) AS `TotalOrders`,
33            AVG(`Purchase Amount (USD)`) AS `AvgSpend`
34    FROM shopping_trends_updated
35    GROUP BY `Location`
36    ORDER BY `AvgSpend` DESC;
37 •  SELECT `Location`, COUNT(*) AS `TotalOrders`,
38            AVG(`Purchase Amount (USD)`) AS `AvgSpend`
39    FROM shopping_trends_updated
40    GROUP BY `Location`
41    ORDER BY `AvgSpend` DESC;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Location	TotalOrders	AvgSpend
Alaska	72	67.5972
Pennsylvania	74	66.5676
Arizona	65	66.5538
West Virginia	81	63.8765
Nevada	87	63.3793
Washington	73	63.3288
North Dakota	83	62.8916
Virginia	77	62.8831
Utah	71	62.5775

Result 5 | Result 6

Task 3: SQL for Data Analysis

5) Create views for analysis

```
43  CREATE VIEW CategorySalesSummary AS
44  SELECT
45      Category,
46      count(*) AS ItemsSold,
47      SUM(`Purchase Amount (USD)`) AS TotalSales
48  FROM shopping_trends_updated
49  group by Category;
50  SELECT * FROM categorysalessummary LIMIT 10;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Conter

Category	ItemsSold	TotalSales
Clothing	1737	104264
Footwear	599	36093
Outerwear	324	18524
Accessories	1240	74200

6) Optimize queries with indexes

```
56  CREATE INDEX idx_gender   ON shopping_trends_updated(`Gender`);
57  CREATE INDEX idx_location ON shopping_trends_updated(`Location`);
58  SHOW INDEXES FROM shopping_trends_updated;
59
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Table	Non_unique	Key_name	Seq_in_index	Column_
shopping_trends_updated	1	idx_gender	1	Gender
shopping_trends_updated	1	idx_location	1	Location

Result Grid | Form Editor

Result 12 | Read Only