

ANSWER QUESTIONS (SQL)

1-Customers

1. How does gender distribution correlate with product purchases?

```
SELECT c.Customer_Gender, COUNT(o.Order_ID) AS Total_Orders, COUNT(DISTINCT
p.Product_ID) AS Unique_Products
FROM Orders o
JOIN Customers c ON o.Customer_ID = c.Customer_ID
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY c.Customer_Gender;
```

1.Orders	2. Products	3. Suppliers	4. Customers
Order ID (INT , PK)	Product ID (char)	Product ID (char)	Customer ID (int, pk)
Product ID (char)	Product Name (vchar)	Supplier ID. (char)	Customer Age Group (char)
Customer ID (INT,FK)	Price (float)		Customer Location (Char)
Shipping Cost (float)	Tax Rate (float)		Customer Gender. (Char)
Shipping Method(char)	Discount (float)		
Pure Profit. (float)	Category (char)		
	Stock Level (float)		
	Return Rate (float)		
	Seasonality (char)		
	Popularity Index.(float)		

2. Are customer age groups more likely to purchase a product from a specific product category than others?

```
SELECT c.Customer_Age_Group, p.Category, COUNT(o.Order_ID) AS Total_Orders
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
JOIN Customers c ON o.Customer_ID = c.Customer_ID
GROUP BY c.Customer_Age_Group, p.Category
ORDER BY c.Customer_Age_Group, Total_Orders DESC;]
```

3. Top 3 product popularity based on location.

```
SELECT c.Customer_Location, p.Product_Name, AVG(p.Popularity_Index) AS Avg_Popularity
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
JOIN Customers c ON o.Customer_ID = c.Customer_ID
GROUP BY c.Customer_Location, p.Product_Name
ORDER BY c.Customer_Location, Avg_Popularity DESC
LIMIT 3;
```

2.Shipping and Logistics

1. What's the most used shipping method for each product category?

```
SELECT p.Category, o.Shipping_Method, COUNT(o.Order_ID) AS NUM_Usage
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY p.Category, o.Shipping_Method
ORDER BY p.Category, Method_Usage DESC;
```

2. Are products with higher return rates associated with a specific shipping method?

```
SELECT o.Shipping_Method, AVG(p.Return_Rate) AS Avg_Return_Rate
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY o.Shipping_Method
ORDER BY Avg_Return_Rate DESC;
```

1.Orders	2. Products	3. Suppliers	4. Customers
Order ID (INT , PK)	Product ID (char)	Product ID (char)	Customer ID (int, pk)
Product ID (char)	Product Name (vchar)	Supplier ID. (char)	Customer Age Group (char)
Customer ID (INT,FK)	Price (float)		Customer Location (Char)
Shipping Cost (float)	Tax Rate (float)		Customer Gender. (Char)
Shipping Method(char)	Discount (float)		
Pure Profit. (float)	Category (char)		
	Stock Level (float)		
	Return Rate (float)		
	Seasonality (char)		
	Popularity Index.(float)		

3. Which regions have the highest 'Shipping Costs?

```
SELECT c.Customer_Location, AVG(o.Shipping_Cost) AS Avg_Shipping_Cost
FROM Orders o
JOIN Customers c ON o.Customer_ID = c.Customer_ID
GROUP BY c.Customer_Location
ORDER BY Avg_Shipping_Cost DESC
LIMIT 10;
```

4. What is the most cost-effective shipping method based on the average 'Shipping Cost'?

```
SELECT o.Shipping_Method, AVG(o.Shipping_Cost) AS Avg_Shipping_Cost
FROM Orders o
GROUP BY o.Shipping_Method
ORDER BY Avg_Shipping_Cost ASC LIMIT 1;
```

3-Product

1. What is the overall distribution of stock levels and the risk of stockouts?

```
SELECT Stock_Level, COUNT(Product_ID) AS Product_Count
FROM Products
GROUP BY Stock_Level
ORDER BY Stock_Level ASC;
```

1.Orders	2. Products	3. Suppliers	4. Customers
Order ID (INT , PK)	Product ID (char)	Product ID (char)	Customer ID (int, pk)
Product ID (char)	Product Name (vchar)	Supplier ID. (char)	Customer Age Group (char)
Customer ID (INT,FK)	Price (float)		Customer Location (Char)
Shipping Cost (float)	Tax Rate (float)		Customer Gender. (Char)
Shipping Method(char)	Discount (float)		
Pure Profit. (float)	Category (char)		
	Stock Level (float)		
	Return Rate (float)		
	Seasonality (char)		
	Popularity Index.(float)		

2. What are the Top 3 suppliers that supply the top 3 profitable products?

```
SELECT p.Product_ID, p.Product_Name, SUM(o.Pure_Profit) AS Total_Profit
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY p.Product_ID, p.Product_Name
ORDER BY Total_Profit DESC LIMIT 3)
SELECT s.Supplier_ID, pp.Product_Name, pp.Total_Profit
FROM Suppliers s
JOIN ProductProfit pp ON s.Product_ID = pp.Product_ID
ORDER BY pp.Total_Profit DESC;
```

Another way

```
SELECT p.Product_ID, p.Product_Name, s.Supplier_ID,
SUM(o.Pure_Profit) AS Total_Profit
FROM Orders o
JOIN Products p ON p.Product_ID = o.Product_ID
JOIN Suppliers s ON p.Product_ID = s.Product_ID
GROUP BY p.Product_ID, p.Product_Name, s.Supplier_ID
ORDER BY Total_Profit DESC LIMIT 3;
```

3.What are each City's Trends By Category

```
SELECT c.Customer Location , p.Category,
COUNT(o.Order_ID) AS Total_Orders
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
JOIN Customers c ON o.Customer_ID = c.Customer_ID
GROUP BY c.Customer Location, p.Category
ORDER BY c.Customer Location, Total_Orders DESC;
```

1.Orders	2. Products	3. Suppliers	4. Customers
Order ID (INT , PK)	Product ID (char)	Product ID (char)	Customer ID (int, pk)
Product ID (char)	Product Name (vchar)	Supplier ID. (char)	Customer Age Group (char)
Customer ID (INT,FK)	Price (float)		Customer Location (Char)
Shipping Cost (float)	Tax Rate (float)		Customer Gender. (Char)
Shipping Method(char)	Discount (float)		
Pure Profit. (float)	Category (char)		
	Stock Level (float)		
	Return Rate (float)		
	Seasonality (char)		
	Popularity Index.(float)		

4. Which products perform best overall in sales and popularity?

```
SELECT p.Category, p.Product_Name,
SUM(o.Pure_Profit) AS Total_Profit,
AVG(p.Popularity_Index) AS Avg_Popularity
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY p.Category, p.Product_Name
ORDER BY Total_Profit DESC, Avg_Popularity DESC
LIMIT 10;
```

5. Which categories have the highest return rate?

```
SELECT p.Product_Name, p.Category,
p.Return_Rate
FROM Products p
ORDER BY p.Return_Rate DESC LIMIT 10;
```

6.What is the Top product in each category?

```
SELECT p.Category, p.Product_Name,
SUM(o.Pure_Profit) AS Total_Profit
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY p.Category, p.Product_Name
)
SELECT Category, Product_Name, MAX(Total_Profit)
AS Top_Product_Profit
FROM CategorySales
GROUP BY Category, Product_Name
ORDER BY Category;
```

4-Sales and Revenue

1.How do discount, and tax rates affect overall revenue and profit margins?

```
SELECT p.Discount, p.Tax_Rate,
SUM(o.Pure_Profit) AS Total_Profit,
SUM((p.Price - (p.Price * p.Discount / 100) +
(p.Price * p.Tax_Rate / 100)) AS Total_Revenue
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY p.Discount, p.Tax_Rate
ORDER BY Total_Revenue DESC;
```

2. How does the Seasonality impact the sales of specific products?

```
SELECT p.Product_Name, o.Seasonality,
count(o.order_id) AS Total_Orders,
AVG(p.Popularity_Index) AS Avg_Popularity
FROM Orders o
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY p.Product_Name, o.Seasonality
ORDER BY o.Seasonality, Total_Sales DESC;
```

1.Orders

Order ID (INT , PK)
Product ID (char)
Customer ID (INT,FK)
Shipping Cost (float)
Shipping Method(char)
Pure Profit. (float)

2. Products

Product ID (char)
Product Name (vchar)
Price (float)
Tax Rate (float)
Discount (float)
Category (char)
Stock Level (float)
Return Rate (float)
Seasonality (char)
Popularity Index.(float)

3. Suppliers

Product ID (char)
Supplier ID. (char)

4. Customers

Customer ID (int, pk)
Customer Age Group (char)
Customer Location (Char)
Customer Gender. (Char)

3.Which customer groups (age, gender) contribute the most to revenue?

```
SELECT c.Customer_Age_Group, c.Customer_Gender,
SUM((p.Price - (p.Price * p.Discount / 100) + (p.Price * p.Tax_Rate / 100)) AS Total_Revenue
FROM customers c
JOIN orders o ON c.Customer_ID = o.Customer_ID
JOIN Products p ON o.Product_ID = p.Product_ID
GROUP BY c.Customer_Age_Group, c.Customer_Gender
ORDER BY Total_Revenue DESC;
```