EDA for Amazon sales analysis

--1. Get the count of orders by their status
SELECT Status, COUNT(*) AS order_count
FROM amazon_sales
GROUP BY Status;

 --2. Get the count of orders by fulfillment type:
SELECT Fulfilment, COUNT(*) AS order_count
FROM amazon_sales
GROUP BY Fulfilment;

 --3.Distribution of Ship Services
SELECT ship_service_level, COUNT(*) AS order_count
FROM amazon_sales
GROUP BY ship_service_level;

--4.Find duplicate Order IDs:
SELECT Order_ID, COUNT(*) AS duplicate_count
FROM amazon_sales
GROUP BY Order_ID
HAVING COUNT(*) > 1;

 --5.Identify the top 10 most ordered SKUs:
SELECT SKU, COUNT(*) AS order_count
FROM amazon_sales
GROUP BY SKU
ORDER BY order_count DESC
LIMIT 10;

--6.Calculate the average order quantity per category:
SELECT Category, AVG(Qty) AS avg_quantity
FROM amazon_sales
GROUP BY Category
ORDER BY avg_quantity DESC;

 --7.Analyze the average quantity ordered by shipping service level:
SELECT ship_service_level, AVG(Qty) AS avg_quantity
FROM amazon_sales
GROUP BY ship_service_level
ORDER BY avg_quantity DESC;

 --8.Count orders by shipping state
SELECT ship_state, COUNT(*) AS order_count

```sql
FROM amazon_sales
GROUP BY ship_state
ORDER BY order_count DESC;
```

--9. Calculate the ratio of shipped to cancelled orders:
```sql
SELECT
    (SELECT COUNT(*) FROM amazon_sales WHERE Status LIKE 'Shipped%')::DECIMAL /
    (SELECT COUNT(*) FROM amazon_sales WHERE Status = 'Cancelled') AS
shipped_to_cancelled_ratio;
```

--10. Find the average size ordered for each category (assuming size is numerical or can be mapped to numerical values):
```sql
WITH size_mapping AS (
    SELECT *,
        CASE
            WHEN Size = 'S' THEN 1
            WHEN Size = 'M' THEN 2
            WHEN Size = 'L' THEN 3
            WHEN Size = 'XL' THEN 4
            WHEN Size = 'XXL' THEN 5
            WHEN Size = 'XXXL' THEN 6
            ELSE NULL
        END AS size_numeric
    FROM amazon_sales
)
SELECT Category, AVG(size_numeric) AS avg_size
FROM size_mapping
GROUP BY Category
ORDER BY avg_size;
```


--11. Identify orders with expedited shipping and fulfilled by Amazon
```sql
SELECT *
FROM amazon_sales
WHERE ship_service_level = 'Expedited' AND Fulfilment = 'Amazon';
```

--12. Identify orders that were shipped but not marked as delivered
```sql
SELECT *
FROM amazon_sales
WHERE Status LIKE 'Shipped%' AND COURIER_STATUS IS NULL;
```

--13.Compare performance of different sales channels within each category
```sql
SELECT Category, Sales_Channel, COUNT(*) AS order_count
FROM amazon_sales
```

```sql
GROUP BY Category, Sales_Channel
ORDER BY Category, order_count DESC;


--14.Analyze which categories are more likely to use expedited shipping
SELECT Category, ship_service_level, COUNT(*) AS order_count
FROM amazon_sales
GROUP BY Category, ship_service_level
ORDER BY Category, order_count DESC;


--15. Examine the distribution of sizes for different fulfillment types
SELECT Fulfilment, Size, COUNT(*) AS order_count
FROM amazon_sales
GROUP BY Fulfilment, Size
ORDER BY Fulfilment, order_count DESC;
```