2024

SUPERSTORE SALES ANALYSIS

SUPERSTORE ANALYSIS – REPORT

Problem Statement

In the retail industry, understanding customer behavior and product trends is important for businesses to optimize operations, increase customer satisfaction, and drive growth and profitability. This dataset focuses on sales transactions from a superstore, providing an opportunity to uncover key issues that may influence business performance and ways to address them.

About Dataset

The dataset contains detailed information about Super store's operations including three main tables i.e. orders, customers and employees. The key attributes of all three tables are as under:-

1. Orders Table
    a. ID (Primary Key).
    b. Order_ID.
    c. Ship_Mode.
    d. Segment.
    e. City.
    f. State.
    g. Postal_Code.
    h. Region.
    i. Product_ID.
    j. Category.
    k. Sub_Category.
    l. Product_Name.
    m. Sales.
    n. Quantity.
    o. Discount.
    p. Profit.
    q. Returned.
    r. Total_cost.

s. price_per_unit.

t. Customer_ID (Foreign Key).

u. Employee_ID (Foreign Key).

v. Order_Date.

w. Ship_Date.

2. Customers Table

a. Customer_ID (Primary key).

b. Customer_Name.

c. Segment.

3. Employees Table

a. Employee_ID (Primary Key).

b. Employee_Name.

c. Gender.

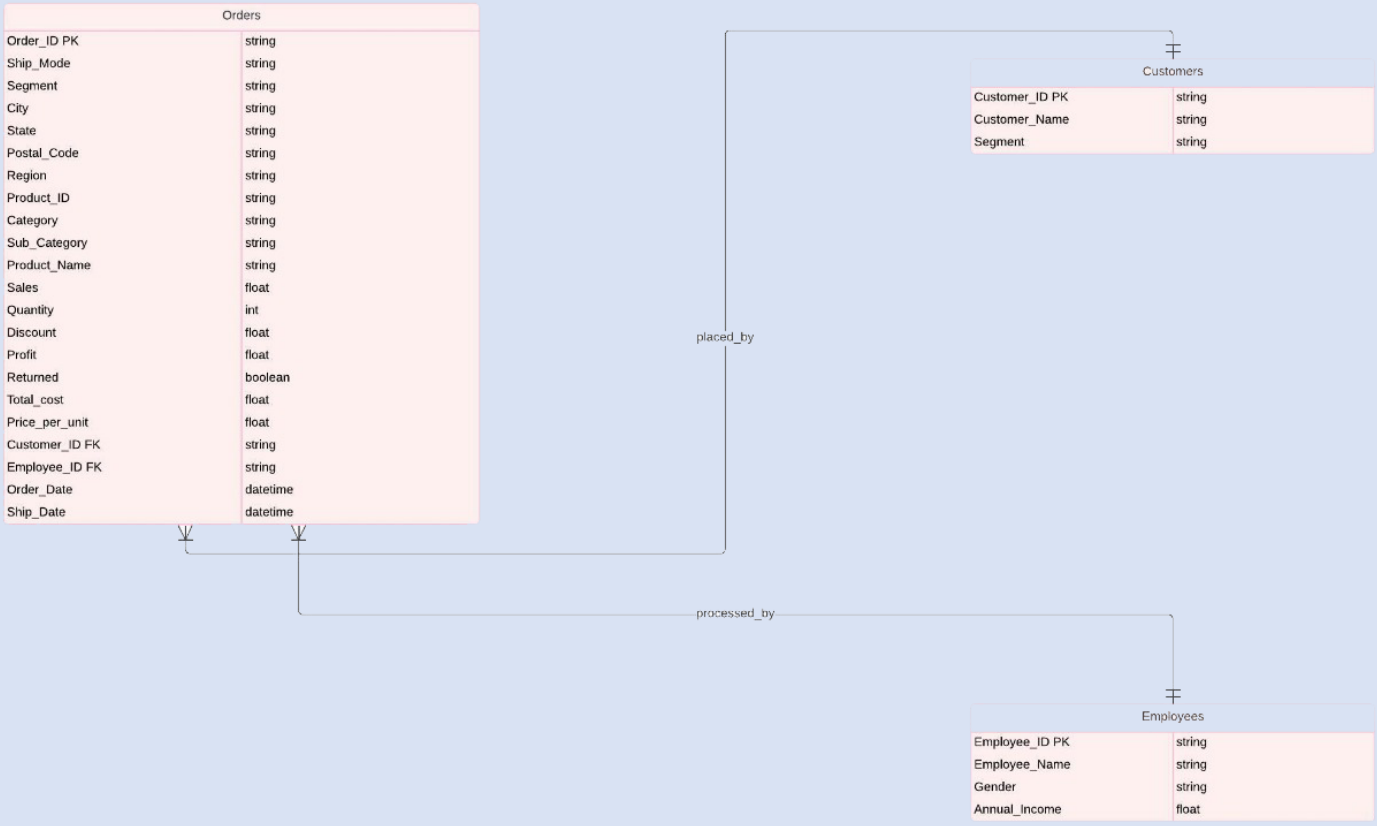
d. Annual_Income.

Data Dictionary

Table	Field Name	Type	Mode	Description
Orders	ID	int	Required	Unique identifier for each order
	Order_ID	varchar(20)	Nullable	Identifier for each order
	Ship_Mode	text	Nullable	Shipping mode of the order
	Segment	text	Nullable	Consumer belongs to which segment of society
	City	text	Nullable	City where the consumer lives
	State	text	Nullable	State where the consumer lives
	Postal_Code	int	Nullable	Postal code where the consumer lives
	Region	text	Nullable	Region where the consumer lives
	Product_ID	text	Nullable	Identifier for each product
	Category	text	Nullable	Product category
	Sub_Category	text	Nullable	Product sub category
	Product_Name	text	Nullable	Product name
	Sales	double	Nullable	Amount of sale generated
	Quantity	int	Nullable	Quantity sold
	Discount	double	Nullable	Discount availed
	Profit	double	Nullable	Profit generated
	Returned	text	Nullable	Orders returned
	total_cost	double	Nullable	Total value of order
	price_per_unit	double	Nullable	Cost per unit
	Customer_ID	varchar(20)	Nullable	Identifier for each customer

	Employee_ID	varchar(20)	Nullable	Identifier for each employee
	Order_Date	date	Nullable	Date on which order was placed
	Ship_Date	date	Nullable	Date on which order was shipped
Customers	Customer_ID	varchar(20)	Required	Unique identifier for each customer
	Customer_Name	text	Nullable	Customer name
	Segment	text	Nullable	Segment to which the customer belongs
Employees	Employee_ID	varchar(20)	Required	Unique identifier for each employee
	Employee_Name	text	Nullable	Employee name
	Gender	text	Nullable	Gender

ERD Diagram



Analytical Questions

Product Analysis

1. **Top Selling Items** – Find the top 5 items with the highest average sales per day?
2. **Product Popularity** – What is region wise the most demanded sub-category?
3. **Profitability Analysis** – Which product, city & region are most & least contributing to total revenue?

Customer Analysis

4. **Customer Demographics** – Give the name of customers who ordered highest and lowest orders from each city?

5. **Customer Segmentation Analysis by State** – Which segment places the highest number of orders from each state?

Shipping Data Analysis

6. **Segment Preference for First-Class Shipping** – Which segment's order is more likely to be shipped via first class?

7. **Shipping Efficiency** – What percentage of total orders were shipped on the same date? What is the average time for orders to get shipped after order is placed varying across different shipping modes?

Sales Analysis

8. **Order Trends Analysis** – How do sales figures vary by month and year, and what are the peak sales periods for the superstore?

9. **Seasonal Demand** – What are the seasonal trends in product sales, and how do specific sub-categories perform during different times of the year?

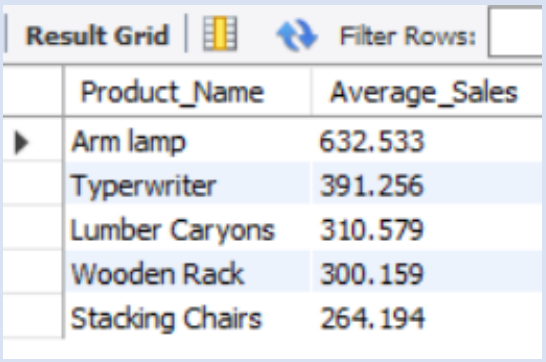
10. **Discount Impact Analysis** - How do discount rates impact the profit margins and sales volume across different product categories?

SQL Queries

Product Analysis

1. **Top Selling Items**

```
SELECT
   Product_Name,
   ROUND(AVG(sales), 3) AS Average_Sales
FROM
   superstore_orders
GROUP BY
   Product_Name
ORDER BY
   Average_Sales DESC
LIMIT 5;
```

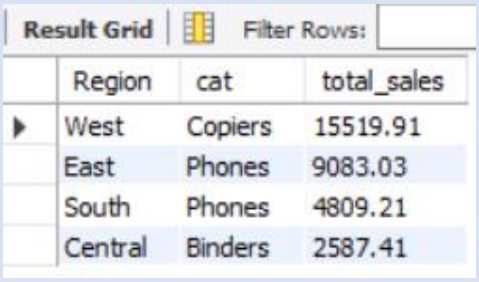


Result Grid | Filter Rows:

	Product_Name	Average_Sales
▶	Arm lamp	632.533
	Typerwriter	391.256
	Lumber Caryons	310.579
	Wooden Rack	300.159
	Stacking Chairs	264.194

2. **Product Popularity**

```
WITH cte as (
SELECT
   Region,
   max(Sub_Category) cat,
   ROUND(SUM(sales), 2) AS total_sales,
   row_number () over (partition by region order by sum(sales) desc) rn
FROM
        superstore_orders
GROUP BY
   Region,
   Sub_Category
ORDER BY
   total_sales DESC)
SELECT
   cte.Region,
   cte.cat,
   cte.total_sales
FROM
   cte
WHERE
   rn = 1;
```

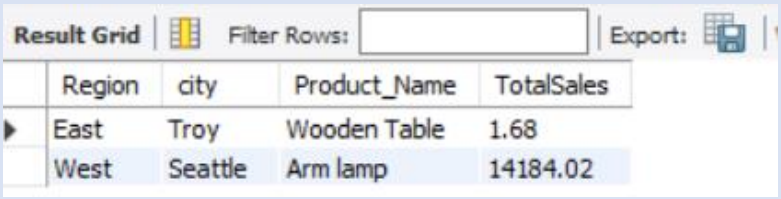


Result Grid | Filter Rows:

Region	cat	total_sales
West	Copiers	15519.91
East	Phones	9083.03
South	Phones	4809.21
Central	Binders	2587.41

3. **Profitability Analysis**

```
(SELECT
    Region,
    city,
    Product_Name,
    ROUND(SUM(sales), 2) AS TotalSales
FROM
    superstore_orders
GROUP BY
    Region, city, Product_Name
ORDER BY
    TotalSales ASC
LIMIT 1)
UNION
(SELECT
    Region,
    city,
    Product_Name,
    ROUND(SUM(sales), 2) AS TotalSales
FROM
    superstore_orders
GROUP BY
    Region, city, Product_Name
ORDER BY
    TotalSales DESC
LIMIT 1);
```



Region	city	Product_Name	TotalSales
East	Troy	Wooden Table	1.68
West	Seattle	Arm lamp	14184.02

Customer Analysis

4. **Customer Demographics**

```
WITH cte AS (
    SELECT
        City,
        ROUND(MAX(sales), 4) AS highest_order,
        ROUND(MIN(sales), 4) AS lowest_order
    FROM
        superstore_orders
    GROUP BY
        City
),
highest_orders AS (
    SELECT
        s.City,
        cte.highest_order,
        cte.lowest_order,
        c.Customer_Name
    FROM
        superstore_orders s
    INNER JOIN
        cte ON s.City = cte.City
          INNER JOIN
        superstore_customers c ON s.Customer_ID = c.Customer_ID
    WHERE
        s.Sales = cte.highest_order
),
```

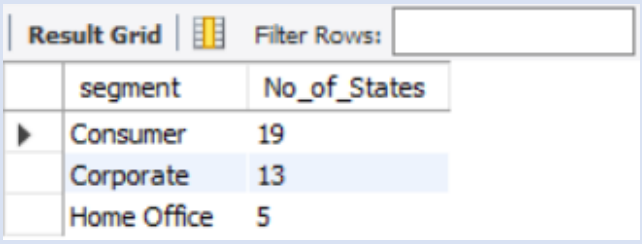
```
lowest_orders AS (
    SELECT
        s.City,
        cte.highest_order,
        cte.lowest_order,
        c.Customer_Name
    FROM
        superstore_orders s
    INNER JOIN
        cte ON s.City = cte.City
    INNER JOIN
        superstore_customers c ON s.Customer_ID = c.Customer_ID
    WHERE
        s.Sales = cte.lowest_order
)
SELECT
    h.City,
    h.highest_order,
    h.Customer_Name AS highest_order_customer,
    l.lowest_order,
    l.Customer_Name AS lowest_order_customer
FROM
    highest_orders h
INNER JOIN
    lowest_orders l ON h.City = l.City
ORDER BY
    h.City;
```

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

City	highest_order	highest_order_customer	lowest_order	lowest_order_customer
Albuquerque	33.29	Angely	23.98	Koray
Amarillo	2453.43	Eloic	263.88	Ainhoa
Anaheim	892.14	Aurelien	5.87	Ariela
Apple Valley	1194.17	Devon	10.56	Victoire
Atlanta	2.78	Perel	2.78	Perel
Aurora	300.77	Hedis	119.8	Bradley
Austin	856.66	Bienne	13.18	Kyliann
Bakersfield	97.84	Adelie	35	Dominic
Bartlett	88.96	Clotilde	88.96	Clotilde

5. **Customer Segmentation Analysis by State**

```
WITH cte AS (
    SELECT
        state,
        segment,
        COUNT(order_id) AS num_orders,
        RANK() OVER (PARTITION BY state ORDER BY COUNT(order_id) DESC)
AS state_rank
    FROM
        superstore_orders
    GROUP BY
        state,
        segment
)
SELECT
    state,
    segment
FROM
    cte
WHERE
    state_rank = 1;
```

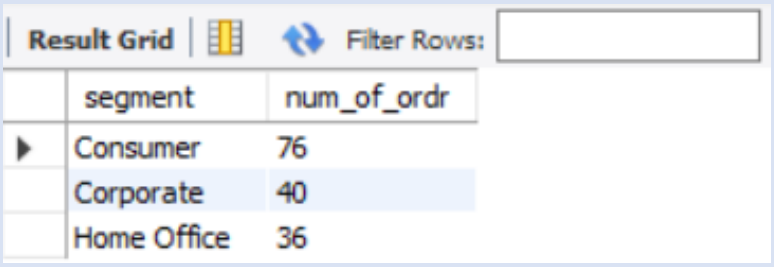


segment	No_of_States
Consumer	19
Corporate	13
Home Office	5

Shipping Data Analysis

6. **Segment Preference for First-Class Shipping**

```
SELECT
   segment,
   COUNT(order_id) AS num_of_ordr
FROM
   superstore_orders
WHERE
   ship_mode = 'First Class'
GROUP BY
   segment
ORDER BY
   num_of_ordr DESC;
```

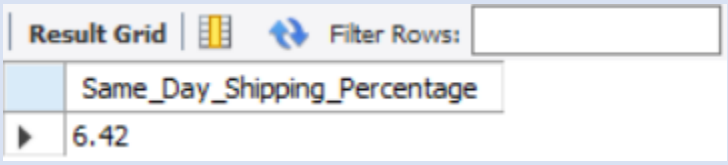


segment	num_of_ordr
Consumer	76
Corporate	40
Home Office	36

7. **Shipping Efficiency**

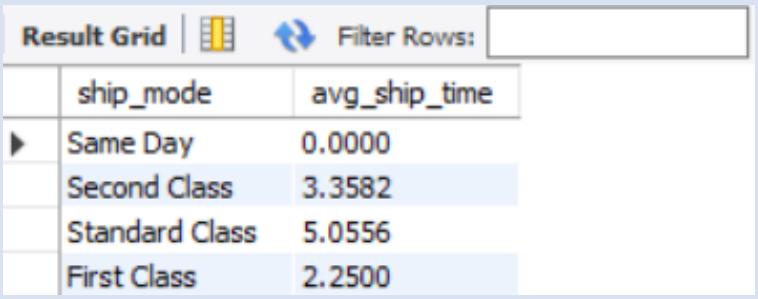
```
SELECT
ROUND((COUNT(DISTINCT Order_ID) / (SELECT COUNT(DISTINCT Order_ID)
AS total_orders FROM superstore_orders)) * 100, 2) AS
Same_Day_Shipping_Percentage
FROM
   superstore_orders
WHERE
```

```
    Order_Date = Ship_Date;
```



Same_Day_Shipping_Percentage
6.42

```
SELECT
    ship_mode,
    AVG(DATEDIFF(ship_date, order_date)) AS avg_ship_time
FROM
    superstore_orders;
GROUP BY
    Ship_mode;
```



ship_mode	avg_ship_time
Same Day	0.0000
Second Class	3.3582
Standard Class	5.0556
First Class	2.2500

Sales Analysis

8. **Order Trends Analysis**

```
WITH cte as (
SELECT
    YEAR(O.Order_Date) AS Year,
    MONTH(O.Order_Date) AS Month,
        ROUND(SUM(O.Sales), 2) AS Total_Sales,
    COUNT(O.Order_ID) AS Number_of_Orders,
    row_number () over (partition by YEAR(Order_Date) order by
COUNT(O.Order_ID) DESC) as rn
FROM
    superstore.superstore_orders as O
GROUP BY
    YEAR(O.Order_Date),
```

```
    MONTH(O.Order_Date)
ORDER BY
    Year,
    Month,
    Number_of_Orders DESC
)
SELECT
        Year, Month, Total_Sales, Number_of_Orders
FROM
        cte
WHERE
        rn<=3
ORDER BY
    Year,
    Number_of_Orders DESC;
```

Result Grid | Filter Rows: | Expor

Year	Month	Total_Sales	Number_of_Orders
2014	9	8183.6	30
2014	12	2869.42	27
2014	8	5112.91	25
2015	12	8830.6	35
2015	11	4967.35	25
2015	10	6126.84	18
2016	12	7109.02	34
2016	10	2806.37	25
2016	5	5820.77	23
2017	9	14108.55	54
2017	12	6829.77	45
2017	8	9081.75	34

9. **Seasonal Demand**

```
WITH cte as (
SELECT
    YEAR(O.Order_Date) AS Year,
    MONTH(O.Order_Date) AS Month,
```

```
    O.Sub_Category,
    Round(SUM(O.Sales), 2) AS Total_Sales,
    SUM(O.Quantity) AS Total_Quantity,
    row_number () over (partition by YEAR(Order_Date) order by SUM(O.Quantity)
DESC) as rn
FROM
    superstore.superstore_orders as O
GROUP BY
    YEAR(O.Order_Date), MONTH(O.Order_Date), O.Sub_Category
ORDER BY
    Year, Month, Sub_Category)
SELECT
        Year,
        Month, Sub_Category, Total_Sales, Total_Quantity
FROM
        cte
WHERE
        rn<=3
ORDER BY
    Year,
    Total_Quantity DESC;
```

Result Grid | Filter Rows: | Export: | Wra

Year	Month	Sub_Category	Total_Sales	Total_Quantity
2014	7	Binders	2237.04	30
2014	11	Paper	505.01	29
2014	9	Paper	306.96	17
2015	12	Binders	584.21	26
2015	12	Furnishings	588.76	18
2015	10	Chairs	3230.84	17
2016	5	Binders	109.8	25
2016	12	Paper	181.29	25
2016	12	Binders	425.59	24
2017	12	Binders	580.82	65
2017	8	Binders	1210.69	33
2017	9	Binders	1090.87	32

10. **Discount Impact Analysis**

```
SELECT
    Category,
    Round(AVG(Discount), 2) AS Avg_Discount,
    Round(SUM(Sales), 2) AS Total_Sales,
    Round(SUM(Profit), 2) AS Total_Profit,
    CASE
        WHEN SUM(Sales) = 0 THEN 0
        ELSE (Profit) / SUM(Sales)
    END AS Profit_Margin
FROM
    superstore.superstore_orders
GROUP BY
    Category
ORDER BY
    Avg_Discount DESC;
```

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

Category	Avg_Discount	Total_Sales	Total_Profit	Discounted_Profit
Technology	0.15	72708.17	13997.38	2099.61
Furniture	0.14	59219.21	2341.16	323.66
Office Supplies	0.14	48576.92	6893.84	986.71

Analysis

1. **Product Analysis**: From 2014 to 2017, the top selling products over all were arm lamp with sale of $632.5, Typewriter with sale of $391.3 and Lumber Caryons with sale of $310.6. Whereas most demanded sub category region wise were Copiers in West, Phones in South and East and Binders in Central. Arm Lamp was the most contributing product to total revenue belonging to Seattle in West region with total sales of $14184.02 and Wooden Table was the least contributing product to total revenue belonging to Troy in East region with total sales of $1.68.
2. **Customer Analysis**: Every city has customers who made the highest and lowest sales. however, following segment of society remained the top buyers in different states. Consumer segment in 19x states, corporate segment in 13x states and home office segment in 5x states.
3. **Shipping Data Analysis**: From 2014 to 2017, different segments preferred shipping their orders by First Class. Consumer segment lead the figures with 76x orders ordered through First Class, whereas, Corporate and Home office followed with 40 and 46 orders through First Class respectively. The overall same day shipping percentage remained at 6.42% of total orders shipped. The average time of shipping across different shipping modes were 2.3 days for First Class, 3.4 days for Second Class and 5.1 days for Standard Class shipping mode.
4. **Sales Analysis**: From 2014 to 2017, sales figure varies from 1 order per month in July, 2015 to 54 order per month in Sept, 2017. The trend shows that the peak sales period usually spans from August to December every year with a few exceptions (May, 2016). The seasonal demand analysis shows that in the later half of the year, Binder and papers are the most in demand items with number of orders leading up to 65 and 29 respectively. Keeping in view the sales and amount of discount given to different category, a significant decrease in profit is seen going up to 80%.

Recommendations for Stakeholders

1. **Product Analysis**: Promotions be provided on top selling products and most demanded sub categories to keep them generating sales. Also to use digital marketing to promote the products not performing well so that a substantial amount of sale is also generated from those products like wooden tables from Troy, East Region.
2. **Customer Analysis**: Customers with highest buys may be offered membership or discounts to make them feel valued. Customers with lowest buys be targeted with niche specific ad campaign to increase sales from them. Home office segment be given priority in ads campaign to generate more sales.
3. **Shipping Data Analysis**: Home office segment be targeted and offered discount through First Class shipping mode to increase their preference for First Class shipping mode. Also, customers ordering through First Class may also be given incentive like discount voucher to encourage standard shipping mode customers to shift to First Class shipping mode.
4. **Sales Analysis**: A well thought out marketing campaign is required to be designed to increase sales in the earlier half of the year. Customer preferences and feedback may be sought to better understand the needs of the customer and offer need specific products to increase sales in the first six months of the year as well.