SUPERSTORE SALES ANALYSIS

Group Members

Mutaher Ijaz

Hareem Farooq

Imran Ahmed

TABLE OF CONTENTS

PROBLEM STATEMENT

PROBLEM STATEMENT

The retail superstore aims to enhance operations and customer satisfaction through data-driven decision-making. They seek to **optimize logistics**, **identify product trends**, and **personalize marketing strategies** by analyzing sales transactions, order processing, and shipping data. Additionally, **evaluating employee sales performances** and **customer behaviors** will enable targeted improvements to drive **growth and profitability** in the competitive retail landscape

UNDERSTANDING THE PROBLEM

CUSTOMER BEHAVIOUR

- Identify customer behaviours
- Identify Seasonal Trends in Sales
- Identify customer area

SALES TRENDS

- Identify most selling items
- Analyze revenue contribution
- Identify Top-performing sub-category

INSIGHTS

- Imporve stocking/ restocking mechanism
- How to engage more customers
- How to increase sales

ABOUT DATASET &
ERD DIAGRAM

ABOUT DATASET

The dataset contains detailed information about Super Store's operations, covering orders, customers, employees, and products

Key Tables and Attributes:

ORDERS

ID, Order_ID, Ship_Mode, Segment, City, State, Postal_Code, Region, Product_ID, Category, Sub_Category, Product_Name, Sales, Quantity, Discount, Profit, Returned, total_cost, price_per_unit, Customer_ID, Employee_ID, Order_Date, Ship_Date

CUSTOMERS

Customer_ID, Customer_Name, Segment

EMPLOYEES

Employee_ID, Employee_Name, Gender, 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

DATA DICTIONARY

Table	Field Name	Type	Mode	Description
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

Table	Field Name	Type	Mode	Description
Employees	Employee_ID	varchar(20)	Required	Unique identifier for each employee
	Employee_Name	text	Nullable	Employee name
	Gender	text	Nullable	Gender
	Annual_Income	int	Nullable	Annual income of the employee

ERD DIAGRAM

Orders	
Order_ID PK	string
Ship_Mode	string
Segment	string
City	string
State	string
Postal_Code	string
Region	string
Product_ID	string
Category	string
Sub_Category	string
Product_Name	string
Sales	float
Quantity	int
Discount	float
Profit	float
Returned	boolean
Total_cost	float
Price_per_unit	float
Customer_ID FK	string
Employee_ID FK	string
Order_Date	datetime
Ship_Date	datetime

Customers	
Customer_ID PK	string
Customer_Name	string
Segment	string

placed_by

processed_by

Employees	
Employee_ID PK	string
Employee_Name	string
Gender	string
Annual_Income	float

ANALYTICAL QUESTIONS

1. **Top 5 items** with the highest average sales per day?

2. What is the region wise **top-performing sub-category** in terms of sales volume & revenue?

3. Which product, city, and region **contribute the most and least** to total revenue?

4. Name of customers who ordered **highest and lowest orders** from each city?

5. Which **segment places the most and largest orders** from each state?

6. Which segment's order is more likely to be **shipped via first class**?

7. What **percentage of orders** were shipped on the **same day**, and what is the **average shipping time** across different modes and product categories?

8. How do **monthly and yearly sales** vary, and what are the peak sales periods for the superstore?

9. What are the **seasonal trends** in product sales & how specific sub-categories perform over the year?

10. How do **discount rates impact profit margins** and sales volume across different product categories?
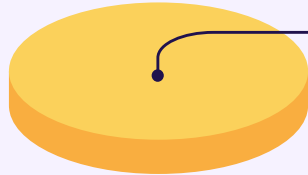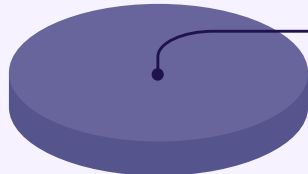
DATA PROCESS
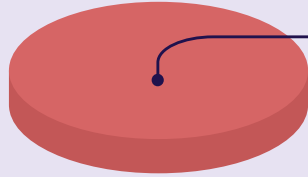
1. **Data Extraction**
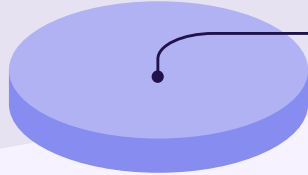
2. **Data Cleaning**

3. **Data Transformation**

4. **Data Validation**

DATA PROCESS

1 **Data Extraction**: Extract Data from Kaggle in three Excel files

2 **Duplicates Removal**: Identified and removed duplicate records

Missing Values Handling: Managed missing data through imputation and deletion

Data Consistency: Standardized formats for dates, names, and categories

3 **Integration:** Merged data from different tables to create a unified dataset

4 **Accuracy Checks**: Verified data accuracy through sample checks & comparisons

Integrity Constraints: Ensured referential integrity across different tables

Sales Dashboard

Total Sales	Total Cost	Total Profits	Total Products Sold
$180.50K	$158.18K	$22.32K	3K

Category Wise Sales

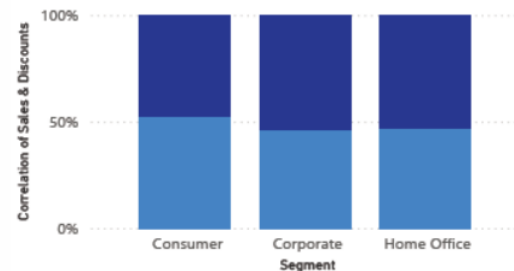● Furniture ● Office Supplies ● Technology

Profit Rolling Average

● Furniture ● Office Supplies ● Technology
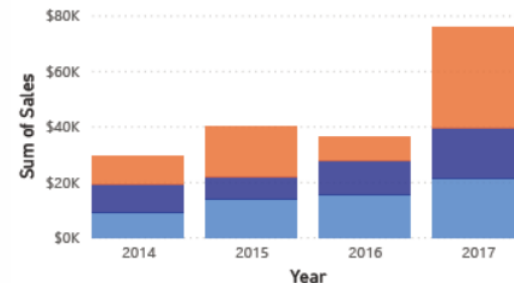
Correlation Between Sales & Discounts

● % Sales ● % Discount

Sales | Customers | Employees

Customers Dashboard
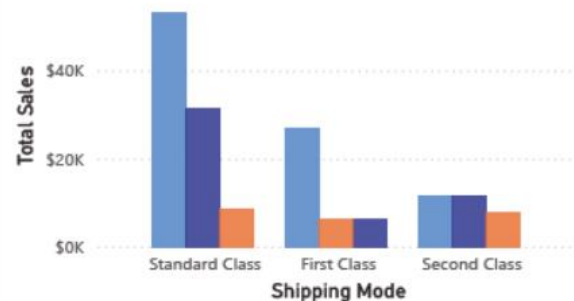
Total Regions
4

Customer Segments
3

Total Orders
800

Average Order Quantity
3.82

## Segment Wise Preference of Shipping Mode

Segment ● Consumer ● Corporate ● Home Office

Total Sales: $40K, $20K, $0K

Shipping Mode: Standard Class, First Class, Second Class

Customer's Order Trend Analysis

(y-axis: 0, 10, 20, 30, 40, 50)

(x-axis: Jan 2014, Jul 2014, Jan 2015, Jul 2015, Jan 2016, Jul 2016, Jan 2017, Jul 2017)

Sales By Segment & Category

Category ● Furniture ● Office Supplies ● Technology

No of Orders: $40K, $20K, $0K

Segment: Consumer, Corporate, Home Office
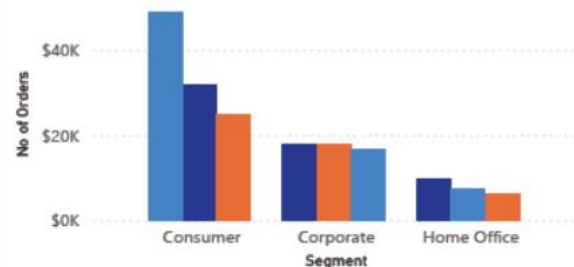
Sales | Customers | Employees

Employees Dashboard

Total Sales
$180.50K

Sales Per Employee
$225.6304
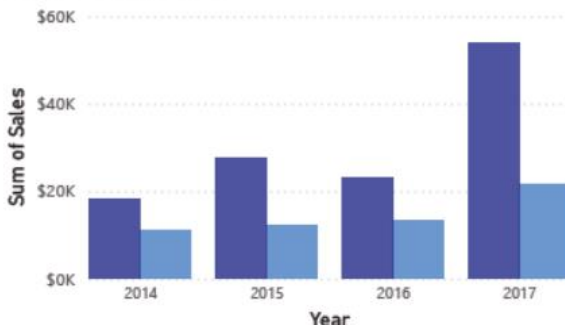
Profit Per Employee
$27.9042

Average Annual Income
53.69K

## Gender Wise Employee Sales Distribution by Year

Gender ● Female ● Male

Region	State	Employee Name	Gender	Year	Month	Quantity	Total Sales	Total Profit	Shipping Mode
West	Washington	Anabella	Male	2017	March	4	$13,999.96	$4,199.99	First Class
East	New York	Maryam	Male	2017	September	11	$2,888.13	$641.81	Standard Class
East	New York	Rowan	Female	2014	September	4	$2,799.96	$839.99	Same Day
East	New York	Kingsley	Male	2015	May	7	$2,799.94	$349.99	Standard Class
West	California	Brooke	Male	2015	December	6	$2,548.56	$318.57	Standard Class
West	California	Heaven	Male	2017	August	9	$2,518.29	$755.49	Standard Class
Central	Texas	Shaindy	Male	2015	October	5	$2,453.43	$0.00	Same Day
East	New York	Mariama	Female	2017	September	5	$2,254.41	$500.98	Standard Class
South	Tennessee	Selma	Male	2017	June	8	$2,239.94	$279.99	Standard Class
Central	Texas	Vincent	Male	2014	July	8	$2,177.58	($5,443.95)	Standard Class
East	Pennsylvania	Sadie	Female	2017	September	7	$1,931.96	($321.99)	Standard Class
West	California	Aidan	Female	2017	November	3	$1,919.98	$240.00	Same Day
West	Arizona	Joshua	Female	2015	August	5	$1,879.96	$235.00	Standard Class
West	Washington	Amy	Male	2016	December	5	$1,747.25	$524.18	Standard Class
West	California	Baptiste	Male	2017	October	14	$1,702.12	$510.64	Standard Class
Total						**3053**	**$180,504.31**	**$22,323.32**	

Top Performing Employees

Region ● East ● West
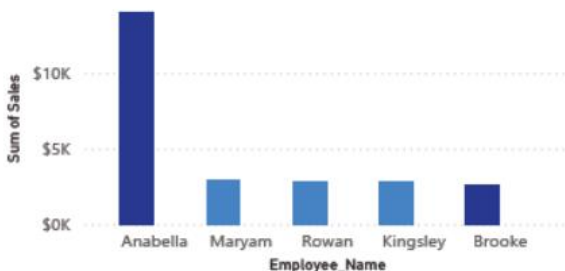
Sales | Customers | Employees

ANALYSIS RESULTS

Product Analysis

Top Selling Products (2014-2017)

Arm Lamp: $632.5

Typewriter: $391.3

Lumber Crayons: $310.6

Most Demanded Sub-Categories (Region Wise)

West: Copiers

South and East: Phones

Central: Binders

Revenue Contributions

Highest: Arm Lamp in Seattle (West) - $14184.02

Lowest: Wooden Table in Troy (East) - $1.68

Customer Analysis

Top Buying Segments by State
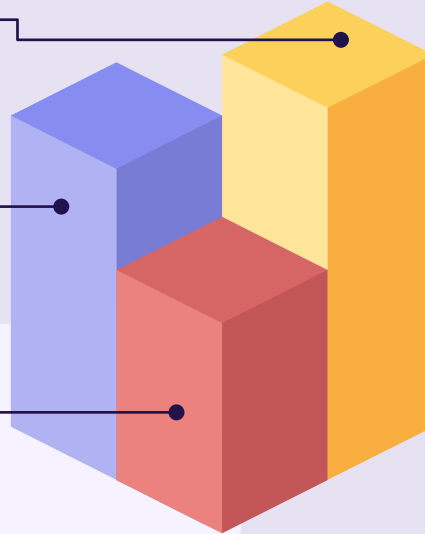
Consumer

Dominates in 19 states

Corporate

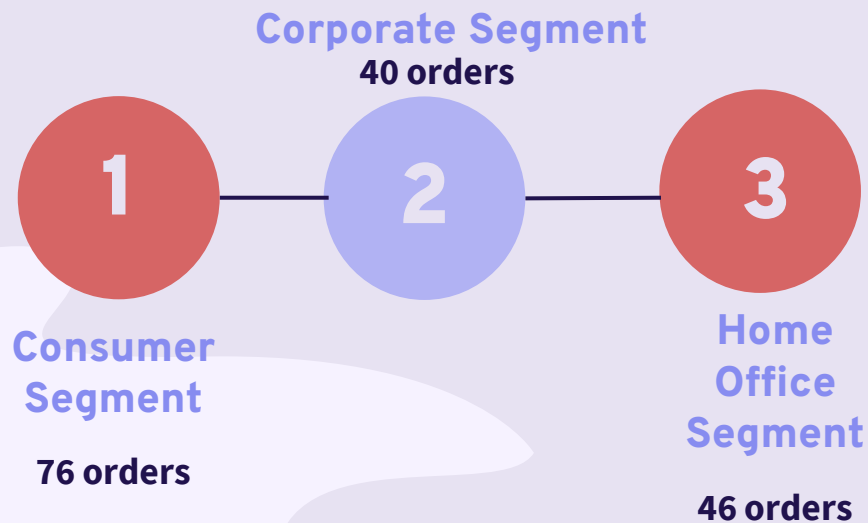Leads in 13 states

Home Office

Tops in 5 states

Sales Variability

Highest and lowest sales recorded in every city

Shipping Data Analysis

Preferred Shipping Methods (2014-2017) is First Class

Corporate Segment
40 orders

1

2

3

Consumer Segment
76 orders

Home Office Segment
46 orders

Same Day Shipping is 6.42% of total orders

Shipping Times

First Class
2.3 days

Second Class
3.4 days

1

2

3

Standard Class
5.1 days
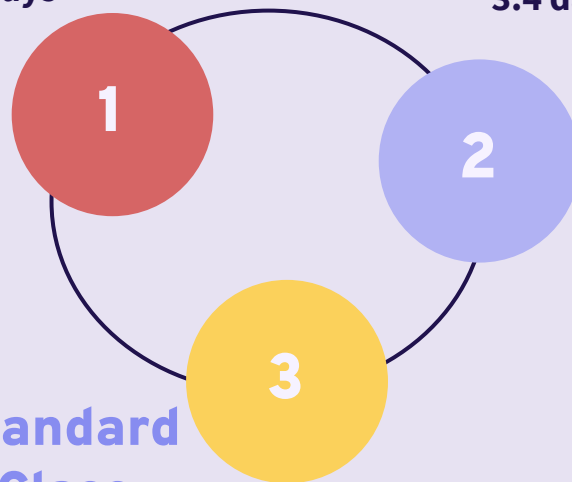
SALES ANALYSIS

Sales Figures (2014-2017)

Low: 1 order in July 2015
High: 54 orders in September 2017

Seasonal Demand

Binders: Up to 65 orders
Papers: Up to 29 orders

Peak Sales Period

August to December

Profit Impact

Significant decrease up to 80% due to discounts

OUR

RECOMMENDATIONS

RECOMMENDATIONS FOR STAKEHOLDERS

Offer Promotions on Top Selling products And Popular Subcategories

Offer Membership or Discounts to High- Buying Customers

RECOMMENDATIONS FOR STAKEHOLDERS

Incentivize First Class Shipping By Providing Discount Vouches

Design a well-Crafted Marketing Campaign that Boost Sales

THANK YOU