

# E-Commerce Sales Data Analysis

*A Case Study*

Tools Used:



# Who's Behind?

Hi, I'm Aritra Biswas,

- Aspiring Data Analyst with 11+ personal projects
- Skilled in SQL, Python, Power BI, and Excel
- Passionate about transforming data into business value

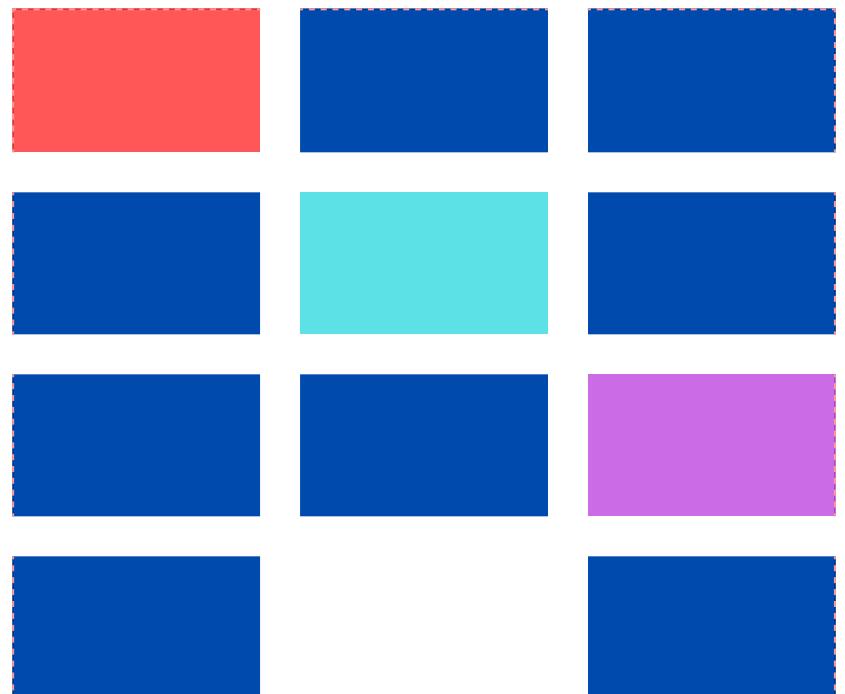


# Project Overview

Dataset: 10,000+ transactions + 793 reviews

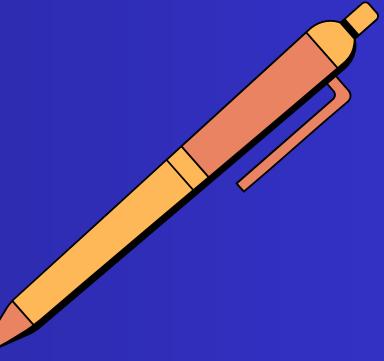
## Objectives:

- a. Improve profitability
- b. Reduce delivery delays
- c. Enhance customer experience

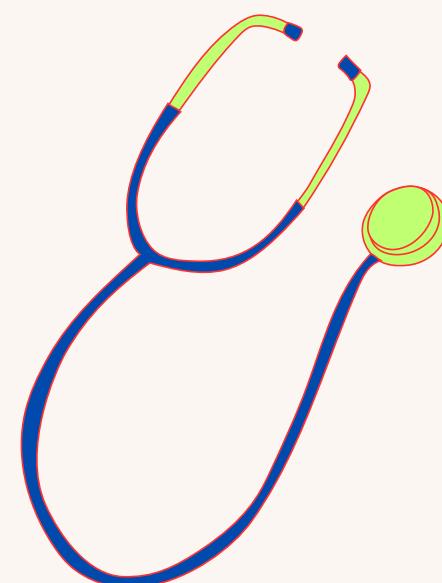
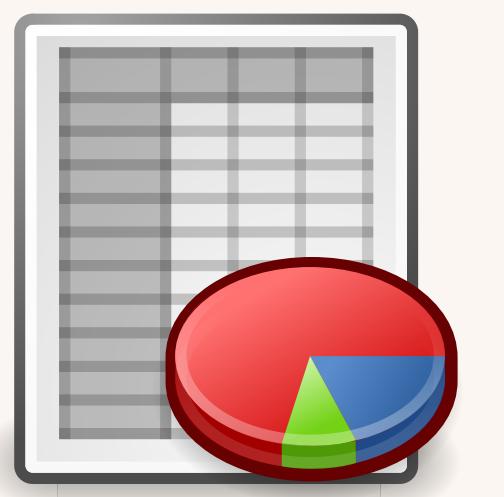
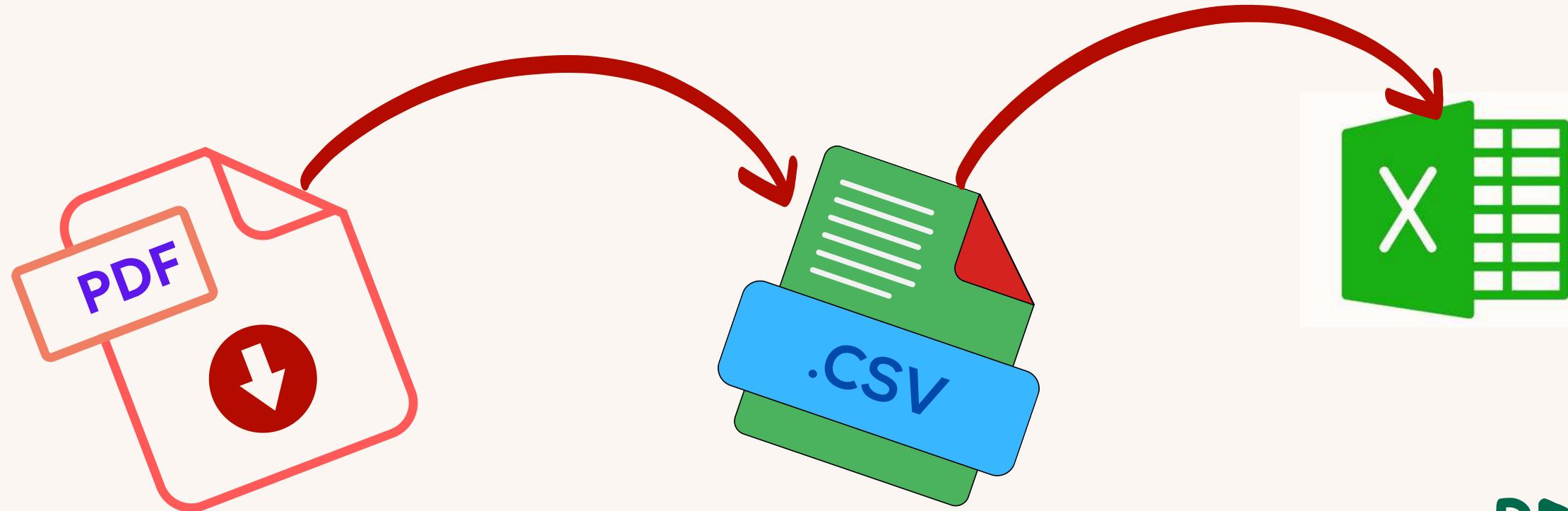


# Business Challenges

1.  Low profitability in Office Supplies
2. Shipping delays in key states 
3.  Poor customer service ratings

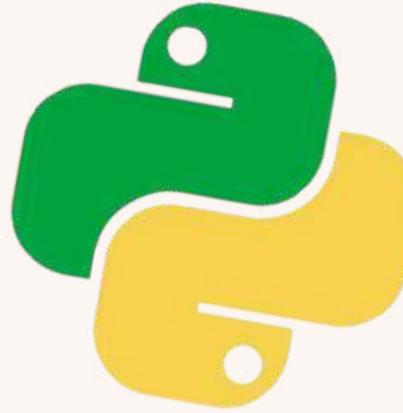


# Understanding the Data



- PDF to CSV Conversion
- Pivot Table Summarization
- Preliminary Data Checks

# Data Preprocessing with



```
df.drop_duplicates(inplace=True) # 216  
df.shape  
  
df.fillna({  
    'sales': df['sales'].mean(),  
    'quantity': df['quantity'].median(),  
    'discount': df['discount'].mean(),  
    'profit': df['profit'].median() }, inplace=True)  
df.isnull().sum()
```

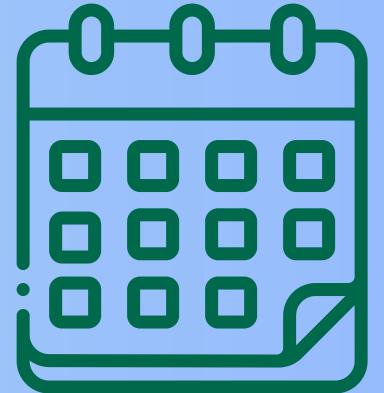


# EDA with

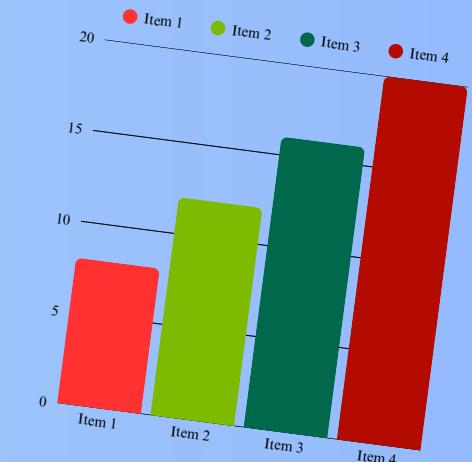


## Monthly Sales Trend !!

```
SELECT ordermonth, COUNT(ordermonth)  
FROM detail  
GROUP BY ordermonth;
```



Insights:



- Highest order months: September, November, December
- Strategic months for campaigns

# EDA with



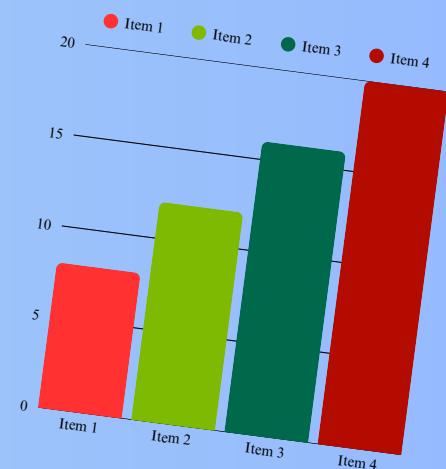
## Profitability Analysis !!

```
SELECT category, COUNT(*)  
FROM detail  
WHERE profit = 0.00  
GROUP BY category;
```



### Insights:

- Office Supplies → Most zero-profit orders
- Deep discounting caused margin erosion



# EDA with



```
SELECT discount, COUNT(*)  
FROM detail  
WHERE profit = 0.00  
AND discount >= 0.5  
GROUP BY discount;
```

## Revenue Leakage



### Findings:

- Discounts over 50% heavily linked to profit loss
- Reevaluate discount policies



# EDA with



## Customer Satisfaction Insights

```
SELECT customerid, customerrating  
FROM review  
WHERE customerrating < 5;
```

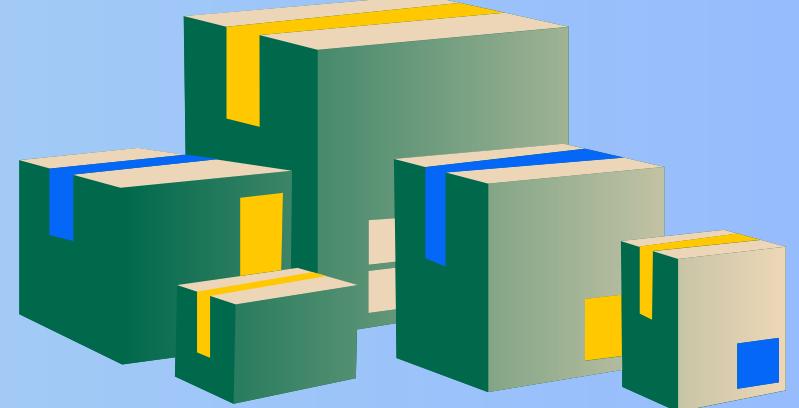


### Insight:

- Customers rated < 5 = high risk for cancellation

### Recommendation:

- Double-confirm orders for these customers



# EDA with

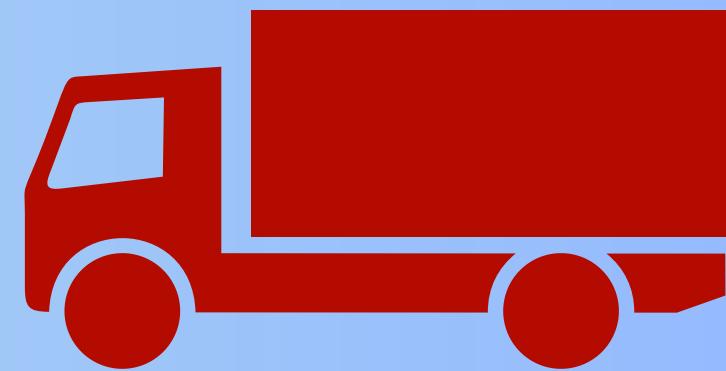
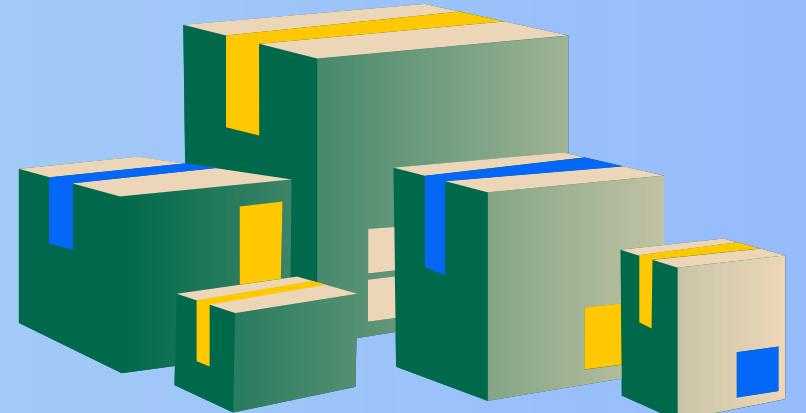


## Shipping Delay Analysis

```
SELECT state, delay, COUNT(*) AS  
delay_count  
FROM detail  
WHERE delay > 5  
GROUP BY state, delay;
```

### Insights:

- Most delayed shipments in: California, New York, Texas
- Logistics improvement needed in these regions

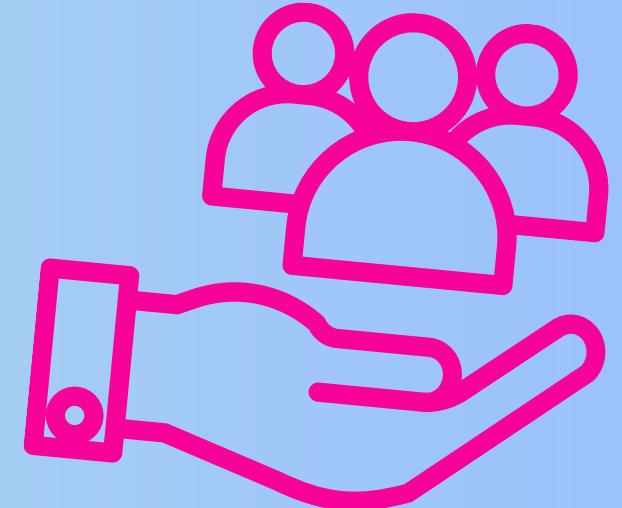


# EDA with Used Case



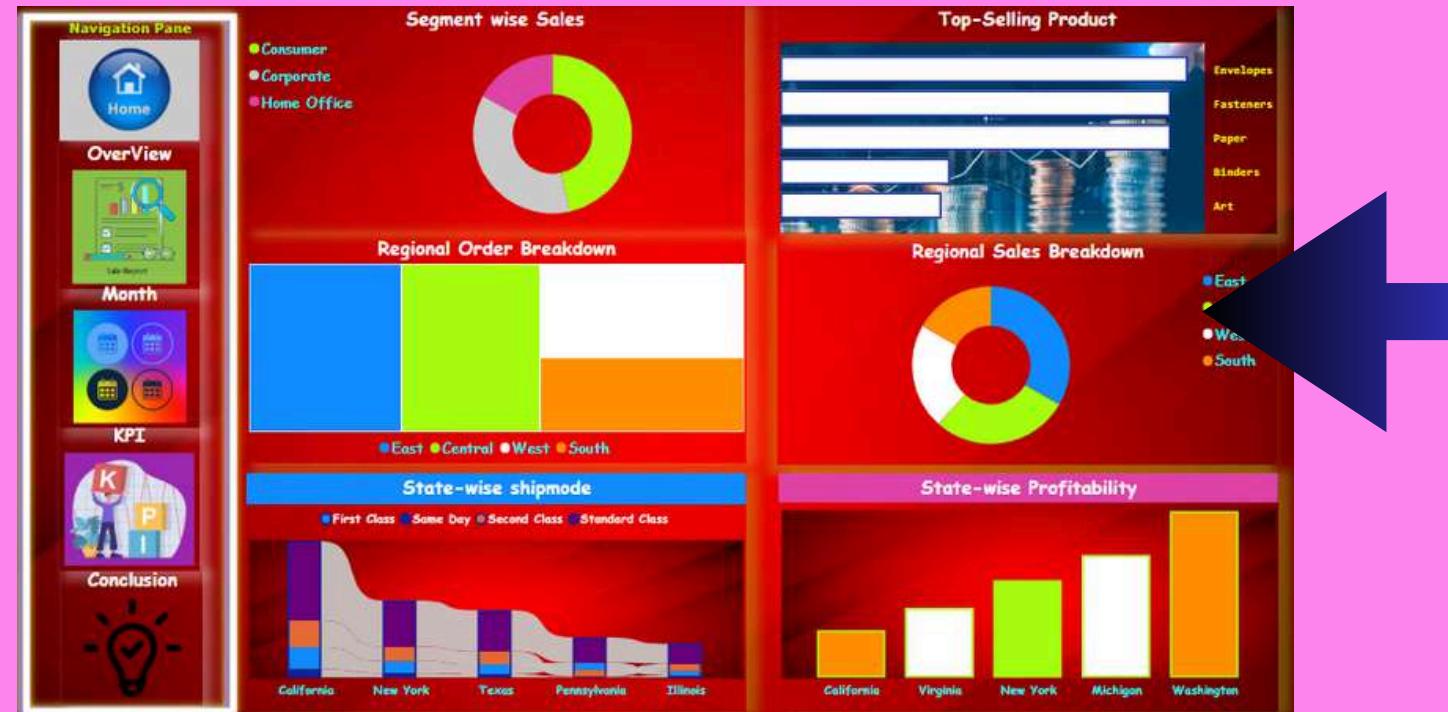
## Customer Segmentation

```
WITH ranked_subcategories AS (
    SELECT segment, subcategory, COUNT(*)
    AS purchases,
        RANK() OVER (PARTITION BY
    segment ORDER BY COUNT(*) DESC) AS rnk
    FROM detail
    GROUP BY segment, subcategory
)
SELECT * FROM ranked_subcategories
WHERE rnk <= 3;
```



- Enables targeted marketing by segment
- Improves upsell and cross-sell success

# Power BI Dashboard





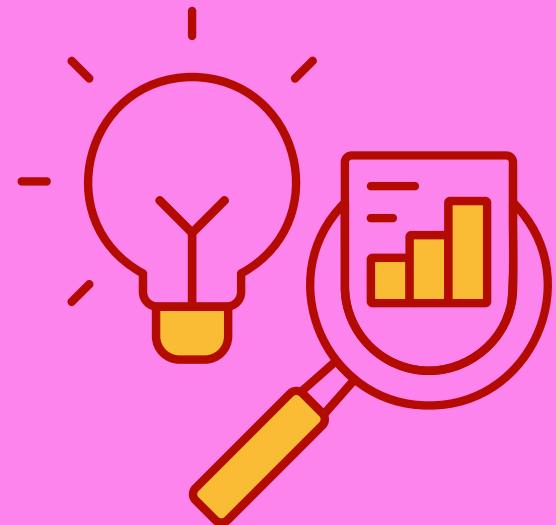
# Power BI Dashboard

To bring analytical insights, I developed an interactive Power BI dashboard designed for clarity, navigation, and stakeholder actionability. The report features a 5-page layout, accessible through a streamlined navigation pane, each focused on distinct layers of business insights.

Insights:

- ₹22 Lakh Total Revenue
- ₹458 Avg. Order Value
- Sales from 531 cities

22 L



## Recommendations:

- Review pricing & discount policies
- Improve logistics in top delay states
- Launch loyalty programs for high-value customers
- Use segment-based promotions for upselling

## Forecasted Impact:

- +11% Profit Margin
- +7–9% Revenue Growth
- +22% Better Delivery Experience

11%



# Thank you very much!

**SUBSCRIBE**

for more projects and insights!



Keep analyzing. Stay curious.

