

# AB4 Sales Analysis Dashboard

---

**Prepared By:** Ankon Banik

**Tools Used:** Power BI, MySQL, Excel

## 1. 🎯 Objective

The purpose of this dashboard is to analyze collected sales data and visualize key performance indicators to support strategic business decisions. The dashboard includes:

- **Top Tiles:**

- Order Quantity
- Order Sales
- Order Revenue
- Average Discount
- Number of Customers
- Average Order Value

- **Visual Analytics:**

- Orders by Category and City
- Profitable Segments
- Preferred Ship Mode
- Profits and Sales by Sub-Category
- Shipment Days by Sub-Category
- Yearly Sales & Profit Trends
- Sales vs. Profit Comparison

- **Interactive Filters:**

- Order Date
- Category
- Segment
- Region
- City

## 2. 📁 Data Collection

The data was sourced from an Excel file provided via a YouTube video description. This dataset contains detailed sales transaction records with the following structure:

- **21 Columns Total:**

- 2 Date columns (Order Date, Ship Date)
- 2 Currency fields (Sales, Profit)
- 2 Numerical values (Quantity, Postal Code)

- 1 Percentage field (Discount)
- 14 Text fields (e.g., Product Name, City, Region, etc.)

### 3. Data Cleaning & Preparation

The raw data was cleaned and processed with the following steps:

1. **Date Formatting:** Converted mm/dd/yyyy to yyyy-mm-dd.

2. **Spelling Correction:** Fixed inconsistencies in City names.

3. **Handling Duplicates:**

- Duplicate Order\_ID and Product\_ID pairs were found with identical values in most columns but varied in Quantity.

- These were treated as valid multi-line items and aggregated to avoid underreporting.

- Aggregation was applied to Quantity, Sales, Discount, and Profit.

Rationale: These rows likely represent either:

- Multiple line items for the same product in a single order
- Partial or split shipments
- Manual reporting entries

4. **Currency Fields:** Converted Sales and Profit fields from currency to numeric formats.

5. **CSV Conversion:** Saved the cleaned Excel data as a .csv file for import into MySQL.

6. **MySQL Integration:** A database named 'sa' and a table 'sales\_order' were created.

### 4. MySQL Table Structure

```
CREATE TABLE sales_order (  
  Row_ID INT PRIMARY KEY,  
  Order_ID VARCHAR(20) NOT NULL,  
  Order_Date DATE NOT NULL,  
  Ship_Date DATE NOT NULL,  
  Ship_Mode VARCHAR(20),  
  Customer_ID VARCHAR(10),  
  Customer_Name VARCHAR(50),  
  Segment VARCHAR(20),  
  Country VARCHAR(50),  
  City VARCHAR(50),  
  State VARCHAR(50),  
  Postal_Code VARCHAR(10),  
  Region VARCHAR(20),  
  Product_ID VARCHAR(20),  
  Category VARCHAR(50),
```

```
Sub_Category VARCHAR(50),
Product_Name VARCHAR(255),
Sales DECIMAL(10, 2),
Quantity INT,
Discount DECIMAL(5, 2),
Profit DECIMAL(10, 2),
UNIQUE (Order_ID, Product_ID),
INDEX idx_order_date (Order_Date),
INDEX idx_customer_id (Customer_ID)
);
```

---

## 5. Dashboard Development in Power BI

- Connected to the MySQL database.
- Built relationships between tables (e.g., Customer, Product, Sales).
- Designed interactive slicers for filter options.
- Used stacked bars, line graphs, and cards for visualization clarity.

## 6. Insights & Observations

- **New York City** has the highest revenue contribution.
- **Segment: Consumer** is the most profitable across all regions.
- **Sub-category: Phones** has high sales but low profit — possibly due to discounting.
- Average shipment days are highest for **sub-category: Art** — opportunity to optimize logistics.

## 7. Conclusion

The dashboard successfully meets the goal of presenting comprehensive sales data with dynamic filters and insightful visualizations. It enables stakeholders to track performance, identify growth areas, and make informed decisions.