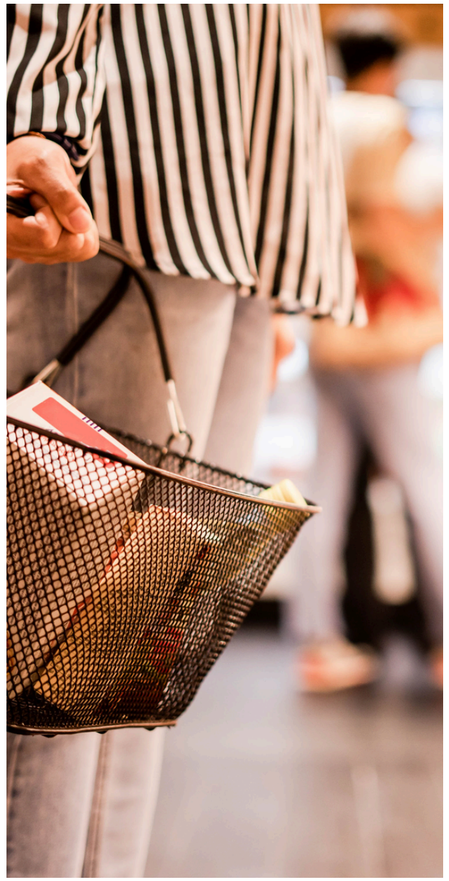
FUNDAMENTAL OF MODERN DATA BIT2053

Audrie Anne Jeffry
Eddy Bin Linus
Grace Micshelle Lynn Majilin
Lyon Carlos Esteban
Stanley Clarence Stefanus



INTRODUCTION

- Decision-making are surrounded by data analytics in this modern business world
- Terabytes records of customer behavior and sales transaction are created by the retail and E-Commerce
- Business Intelligence (BI) tools have been used by many organization to process data to transform it into useful information for future business improvement

BUSSINES SCENARIO

SCENARIO

- Global Superstore is a multinational retail company that operates in various regions around the world
- Companies faces challenges in thier analysis which is give problem for decision-making

QUESTION

- Which regions are most profitable?
- Which product contribute the most to profitability?
- How different customer segments contribute to revenue?

PROBLEM

- Superstore dataset is unclean
- Unclear profit & sales trends
- Need BI tools for data transformation and insights.

DATA PREPARATION & SCHEMA

DATASET DESCRIPTION

Dataset:

- Global Super Store (Kaggle, 2011–2014) – 25,000+ transactions from US, APAC, EMEA, LATAM, Africa, and Canada.

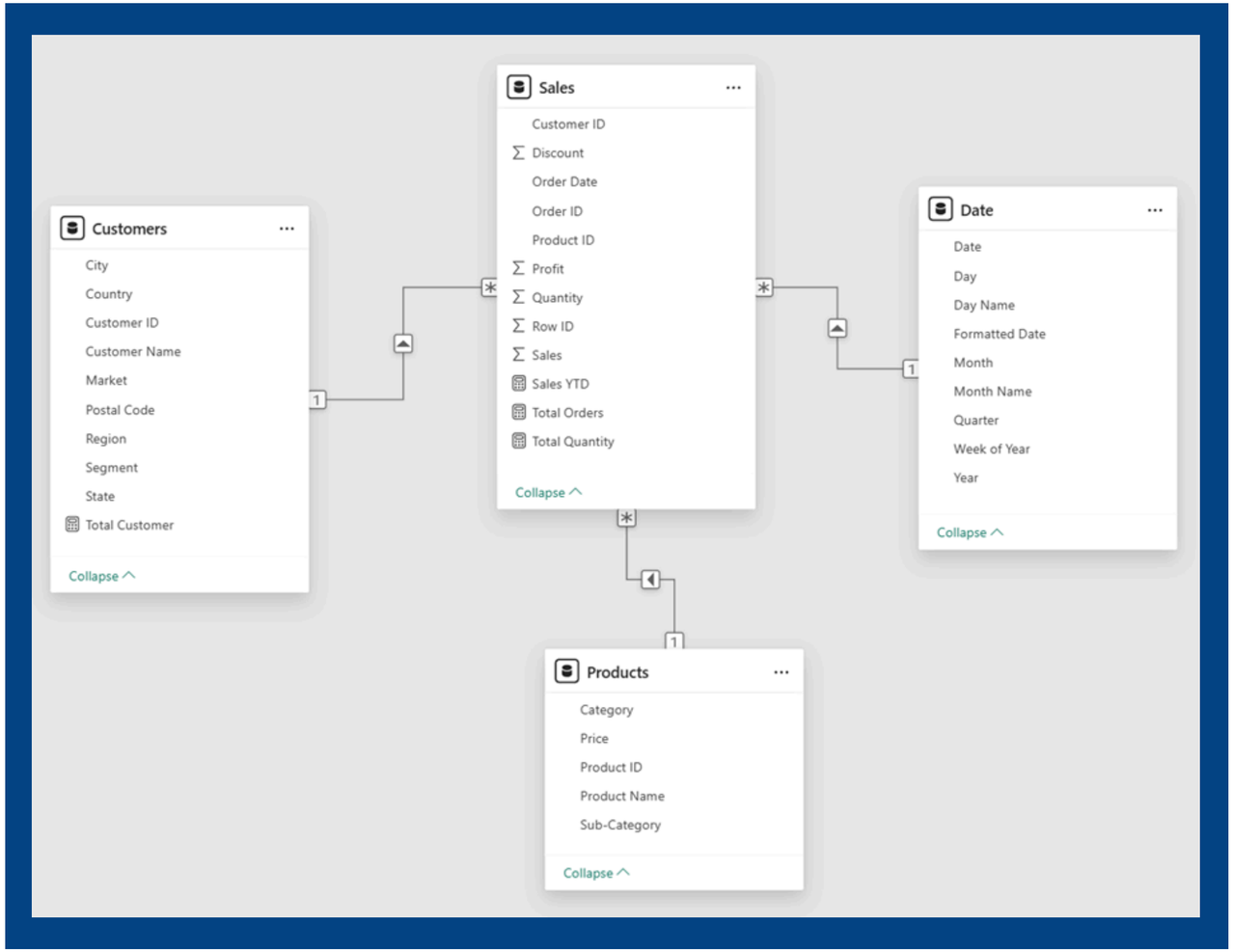
Entities:

- **Customers:** Customer ID, Name, Segment, and Location
- **Products:** Product ID, Name, Category, and Sub-Category.
- **Sales:** Transactional data with order ID, order date, sales, quantity, discount, and profit.
- **Date:** A generated time dimension table covering the dataset range.

DATA CLEANING AND PREPROCESSING

- **Remove Duplicate:** remove a duplication or redundant data in dataset
- **Data Validation:** checking existing data to ensure its quality and values
- **Change Format:** change data into more suiable and consistent data

DATA MODEL
(STAR SCHEMA DESIGN)



- **Fact Table:** Sales (Order ID, Sales Amount, Quantity, Discount, Profit).

- **Dimension Tables:** Customers, Products, Date.

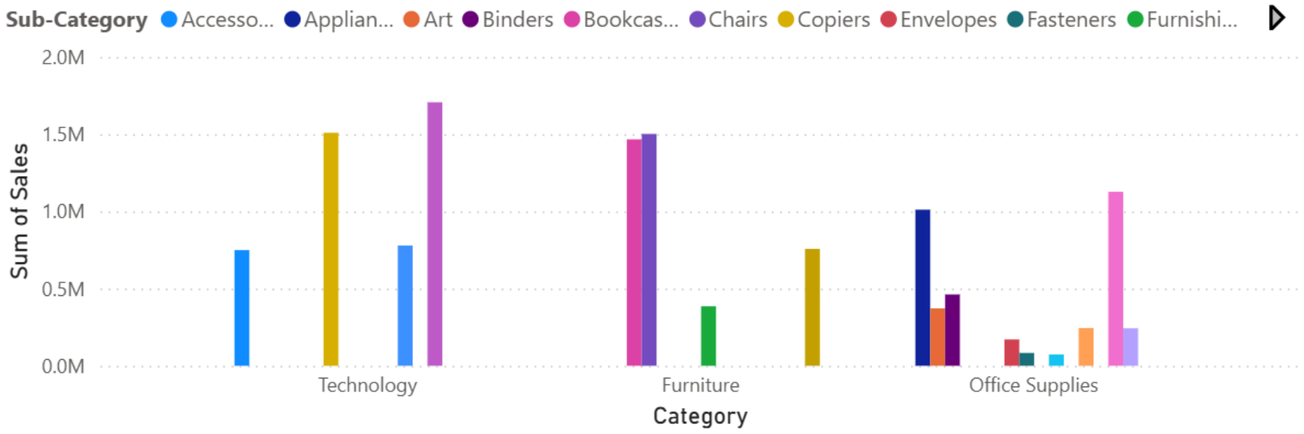
Relationships:

- Sales → Customers (Many-to-One).
- Sales → Products (Many-to-One).
- Sales → Date (Many-to-One).

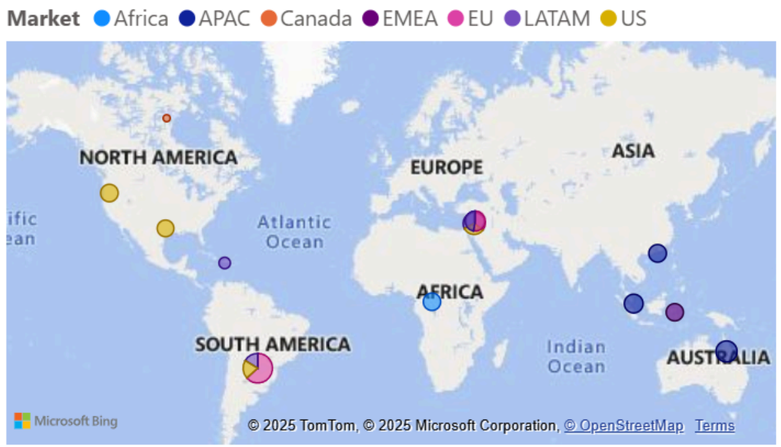
DASHBOARD

12.64M	1590	25K	178K
Sum of Sales	Total Customer	Total Orders	Total Quantity

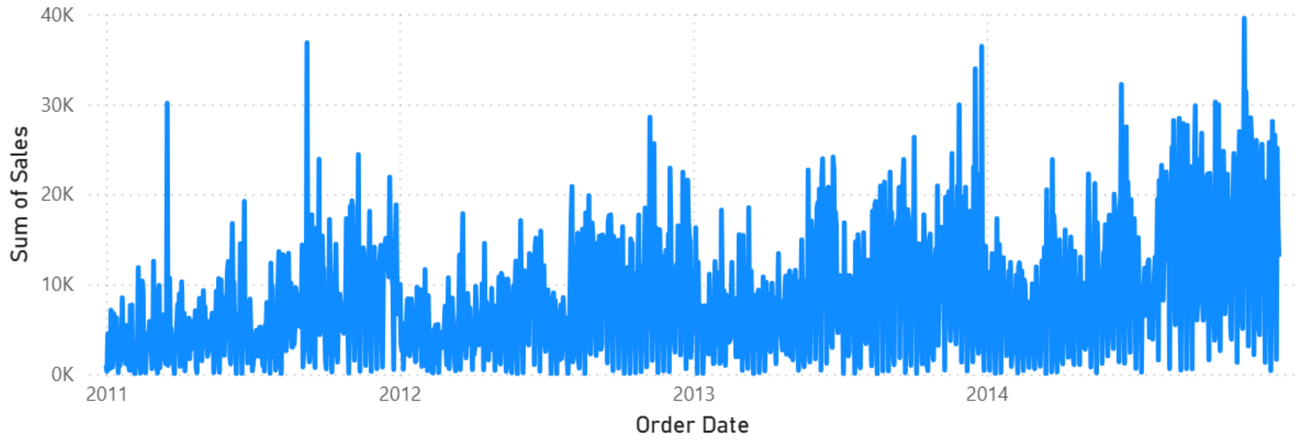
Sum of Sales by Category and Sub-Category



Sum of Sales by Region and Market



Sum of Sales by Order Date



First Category by Order Date

Saturday, 1 Janu...	Sunday, 2 Januar...	Monday, 3 Januar...
Tuesday, 4 Janua...	Wednesday, 5 Ja...	Thursday, 6 Janu...
Friday, 7 January,...	Saturday, 8 Janu...	Sunday, 9 Januar...

KEY INSIGHTS

KPI CARDS

- Total Sales: **12.64M** (strong overall revenue)
- Customers: **1,590** (high repeat purchase rate)
- Orders: **25K** (steady order flow)
- Quality Sold: **178K** (average unit price ≈ 71)

SALES BY CATEGORY & SUB-CATEGORY

- Technology leads revenue: Copiers & Accessories highest performing.
- Furniture stable but less profitable.
- Office Supplies lower sales → potential for promotions.

REGIONAL SALES MAP

- **North America** dominates (US-driven).
- **APAC** emerging growth potential.
- **EMEA** stable but underperforming compared to US.

SALES TREND

- **Peaks:** late 2012 & mid-2014 → linked to seasonal promotions
- Strong **upward trend**, expanding market demand.

SUMMARY

- Analyzed Superstore dataset using Power BI
- Cleaned & modeled sales, profit, and discount data
- Built an interactive dashboard with filters & drill-downs
- Seasonal campaigns correlate with sales peaks.

RECOMMEDATIONS

- Strengthen Corporate Segment Engagement
- Expand the Technology Category
- Enhance Furniture Profitability
- Target Emerging APAC Markets

CONCLUSION

- Project demonstrates use of BI in modern data
- Power BI enabled data-driven decision making, turning raw data into actionable information
- Future work: integrate real-time data, analyse shipping costs, and add predictive forecasting.

THANK YOU