**Superstore Sales Analysis using Power BI**

**1. Project Title**

**Superstore Sales**

**2. Objective**

* Identify the regions with the highest and lowest total sales and profits.
* Analyze top and bottom-performing product categories and sub-categories.
* Understand customer segments and their contribution to revenue.
* Evaluate the impact of shipping modes on business performance.

**3. Dataset Description**

The dataset used is Superstore Sales Dataset.csv, containing 9,800 records of sales transactions. It includes information on order details, customers, geography, product lines, and revenue.

**Key Fields:**

* **Order ID**: Unique identifier for each transaction
* **Order Date / Ship Date**: Timeline of transactions
* **Customer Name / Segment**: Buyer details
* **Region / State / City**: Geographic coverage
* **Category / Sub-Category / Product Name**: Product classifications
* **Sales**: Revenue from the order
* **Postal Code**: Delivery location detail

**4. Data Preparation**

**Data Modeling  
Model Type:** Star Schema (Flat Table Model)

**Structure:**

* **A single fact table Sales containing:**
  + Transactional details (Order ID, Dates, Customer, Product, Region, etc.)
  + Measures like Sales
* Supporting calculated columns and measures using DAX ( Total Sales, Total Orders, etc.)

**Key Relationships:**

* Date Table [Date] → Sales [Order Date]
* Product Table [Product ID] → Sales [Product ID]
* Customer Table [Customer ID] → Sales [Customer ID]

**Reasoning:**Flat model simplifies visual analysis; DAX allows slicing by Category, Region, Segment, and Time for quick insights.

**Data Cleaning**

* Removed redundant columns.
* Transformed columns (e.g., Order Date → Year-Month hierarchy).
* Created calculated columns/measures (e.g., Yearly Sales Growth)

Data preparation was conducted using Power BI:

* Making calendar table.
* Ensured categorical fields (e.g., Region, Segment, Ship Mode) had no typos or inconsistencies.
* Filtered out unnecessary columns such as Country (all rows have seme data ‘United States’) for cleaner modeling.
* Handled null values in **Postal Code** without imputation due to minimal impact.

**5. Data Analysis Performed**

**Questions**

**Sales Analysis**

**1. What were the totals when broken out by product category?**

* **Technology:** Highest total sales ~37%
* **Furniture:** Mid-range ~32%
* **Office Supplies:** Typically, lowest sales share ~31%

**2. What are the total sales for each region?**

* **West:** Highest total sales ~32%
* **East:** Second highest ~28%
* **Central:** Third ~23.5%
* **South:** Lowest total sales region ~16.5%

**Customer Performance**

**3. Order Value by Segment**

* **corporate:** ~ $2.5k
* **Home Office:** ~$1.5k
* **Consumer:** ~$0.6k

**4. Who are the top 5 profitable customers?**

* *Sean Miller, Tamara Chand, Raymond Buc, Tom Ashbrook and Adrian Bardon*

**5. Which segment generates the most revenue?**

* Consumer segment leads in total sales volume

**Analysis of Product**

**6. Top 5 profitable products?**

* Canon Copier, video Conferencing, Electric Punch Plastic Comp Binding Machine with Manual Bind, Chairs for Big and Tall and Electric Binding Systems

**7. Which sub-category has the highest sales? Which was the least profitable?**

* **Highest Sales Subcategory:** *Phones, Chairs, and storage*
* **Least Profitable Subcategory:** *Fasteners*

**8. What are the detailed profit and sales figures within the least profitable category?**

* Office Supplies:
  + **Sales:** $705,422
  + Contain the least 10 profitable products

**9. High-ordered low profitable product**

* Paper (Office Supplies)

**Shipping Analysis**

**10. Which shipping mode is used the most?**

* **Standard Class** is the most common ~60% of orders and ~59% of total sales

**11. Share each Ship Mode**

* **Standard:** ~59%
* **Second Class:** ~20%
* **First Class:** ~15%
* **Same Day:** ~6%

**Time Distribution Analysis**

**12. What is the least year?**

* The least year in sales is 2016 although the number of orders is higher than 2015 but ~52% of them from office supplies that have low price and least year in orders is 2015.

**13. How are sales distributed throughout the year?**

* **Peak Months:** November and December

**14. What are the best-selling months, quarters, and common days every month?**

* Remarkable annual growth in all regions, especially inQ4 (Oct–Dec)
* **Common High-Sales Days:** Mondays (holiday sales spike)

**Analysis of Cities and States**

**15. What are the most profitable cities or states?**

* **States:** California, New York (which represent ~ 48% of total sales) depending on the large amount of orders
* **Cities:** New York City, Los Angeles and Seattle

**16. Is there a concentration of sales in certain geographic areas?**

* Yes — the **West Coast and East** dominate in sales volume and profit and orders

**17. Top Selling Category by Region**

* Office Supplies and Furniture in the Weast
* Technology in the East

**DAX Queries:**

Calendar = ADDCOLUMNS (

    CALENDAR (DATE (2015,01,04), DATE (2018,12,30)),

    "Year”, YEAR([Date]),

    "Month”, MONTH([Date]),

    "Day”, DAY([Date]),

    "Quarter", "Q" & QUARTER([Date]),

    "Weekday", WEEKDAY(Date]),

    "Week Number", WEEKNUM([Date]),

    "Month Name", FORMAT([Date],"MMMM"),

    "Weekday Name", FORMAT([Date],"DDDD"),

    "Start Of Month", DATE(YEAR(Date]), MONTH([Date]), 1)

    )

Pareto Sales % =

IF (

    ISINSCOPE (Location [State]),

    VAR \_\_All Selected States = ALLSELECTED (Location [State])

    VAR \_\_State Sales Table = ADDCOLUMNS (\_\_All selected States, "@Sales”, [Total Sales])

    VAR \_\_Current State Sales = [Total Sales]

    VAR \_\_Cum State Sales Table = FILTER (\_\_State Sales Table, [@Sales] >= \_\_Current State Sales)

    VAR \_\_Cum State Sales = SUMX (\_\_Cum State Sales Table, [@Sales])

    VAR \_\_Total\_Selected\_Sales = CALCULATE ([Total Sales], \_\_All Selected States)

    VAR \_\_Result = DIVIDE (\_\_Cum State Sales, \_\_Total\_Selected\_Sales)

    RETURN

        \_\_Result

)

Total Sales = SUM (Sales [Sales])

Total Customer = DISTINCTCOUNT (Customer [Customer ID])

Total Orders = DISTINCTCOUNT ('Sales'[Order ID])

* Analyzed total sales and profit over time.
* Compared sales performance across all four US regions.
* Ranked categories and sub-categories by revenue and profit.
* Measured shipping performance based on delivery time and method.
* Evaluated segment behavior and contribution to total sales.
* Used slicers and interactive visuals for more capability.

**6. Visualization Steps**

**Viewing Dataset Structure**

Power BI's data view and model view were used to inspect data types and relationships.

**Visualizations Created**

* Bar Charts: Sales by Region, Sales by Category
* Line Charts: Monthly Sales Trends
* Tree Maps: Sales by Sub-Category
* Pie Charts: Segment-wise contribution
* Map Visual: State-wise Sales Distribution

**7. insights & Recommendations**

* **Top Region:** The **West** leads in total sales and profit.
* **Lowest Sales Region:** The **Central** region shows the lowest performance and needs a reinforcement strategy.
* **High-Performing Categories:** Technology generates high profits, especially sub-categories like **Copiers** and **Phones**.
* **Low-Performing Sub-Categories:** **Tables** consistently yield losses due to discounts or high shipping cost.
* **Customer Segments:** **Consumer** dominates in volume, while **Corporate** has higher average transaction value.
* **Shipping Efficiency:** **Same Day** shipping is rarely used but is associated with high-value transactions. (Optimization is recommended).

**8. Challenges Faced**

* Dealing with minor missing values in Postal Codes.
* Creating calculated columns in DAX for dynamic measures.
* Structuring dashboards for both summary and deep-dive analysis without clutter.
* Aligning visuals with stakeholder questions and business goals.

**9. Conclusion**

The Power BI analysis provided comprehensive insights into the superstore’s operations across products, regions, and customer segments. These findings can guide:

* Inventory management for loss-making products
* Strategic focus on underperforming regions
* Enhanced customer segmentation and targeting
* Shipping optimization for profitability