Comprehensive Data Analysis and Sales Forecasting Report

DR. ARPIT YADAV **20/11/2024**

Bhaskar Sharma
2023JULB01197 | JAGSOM

Table of Contents

1. Objective

To contribute to the strategic growth of a business by leveraging **data analysis techniques** and implementing **time series forecasting** to provide accurate sales predictions, actionable insights, and decision-supporting dashboards.

Key Deliverables:

1. Interactive dashboards enabling stakeholders to explore data at various levels.
2. Visualization of trends to assess sales performance and customer behavior.
3. Sales forecasting using time series models for the next 15 days to improve inventory and operational efficiency.
4. Actionable recommendations to enhance sales strategies, optimize planning, and improve customer satisfaction.

---

2. Problem Statement

Despite the growing scale of operations in supermarkets, inefficiencies in data utilization led to challenges in:

1. Accurate sales forecasting.
2. Understanding sales trends across categories and regions.
3. Managing inventory efficiently.

This analysis aims to solve these challenges by identifying KPIs, analyzing sales patterns, and providing a reliable forecasting mechanism.

Key Challenges:

- Inconsistent regional sales trends.
- High return rates for certain product categories.
- Unoptimized delivery systems causing delays.
- Lack of actionable insights from existing sales data.

Goals:

1. Design a time series analysis-based sales forecasting model.
2. Build an intuitive and interactive dashboard for stakeholders.
3. Provide detailed insights and recommendations to improve decision-making.

---

3. Data Requirement

To perform the analysis, the following columns are required:

Field Name	Description
Order ID	Unique identifier for each transaction.
Order Date	Date the order was placed.
Ship Date	Date the order was shipped.
Customer Details	Customer ID, Name, and Segment (Corporate, Consumer, or Home Office).
Product Details	Product Name, Category, Sub-Category, and Quantity.
Sales Details	Sales Amount, Quantity Sold, Profit Margin.
Returns	Indicates whether the product was returned.
Region/Location	State, City, and Region of the customer.
Payment Mode	Payment type used by the customer.

---

4. Data Collection

The dataset used for this analysis is sourced from **Kaggle's E-commerce dataset**:
Dataset Link

Sources:

1. **Sales Data**: Transaction details, customer data, and product information from internal records.
2. **Logistics and Fulfillment**: Data on shipping timelines, order delays, and fulfillment methods.
3. **Promotional Data**: Insights into promotional campaign impacts on sales.

Sample Data:

Order ID	Order Date	Customer Name	Category	Sub-Category	Sales	Profit
CA-2019-160304	01/01/2019	Brendan Murry	Furniture	Bookcases	$73.94	$28.27

---

5. Data Validation

Objective: Ensure data accuracy, consistency, and readiness for analysis.

1. **Completeness Check**:
    - Validate that all critical fields (Order ID, Sales, Profit, etc.) are populated.
2. **Consistency Check**:
    - Uniform formats for dates, states, and numeric values.
3. **Duplicate Records**:
    - Identify and remove any redundant data entries.
4. **Range Validation**:
    - Ensure logical values (e.g., no negative sales, valid product categories).
5. **Missing Data Handling**:
    - Use imputation methods to fill in missing data or exclude irrelevant entries.

---

6. Data Cleaning

Steps:

1. **Manage Missing Data**: Replace missing values for non-critical fields or drop irrelevant rows.
2. **Standardize Formats**: Ensure uniformity in dates, text fields, and numeric precision.
3. **Remove Outliers**: Identify and address anomalies in sales, quantities, or profits.
4. **Consolidate Data**: Combine related categories or merge redundant fields for analysis simplicity.

---

7. Tools and Methodology

Tools Used:

1. **Power BI**: For creating dynamic and interactive dashboards.
2. **Microsoft Excel**: For initial cleaning and validation tasks.

Methodology:

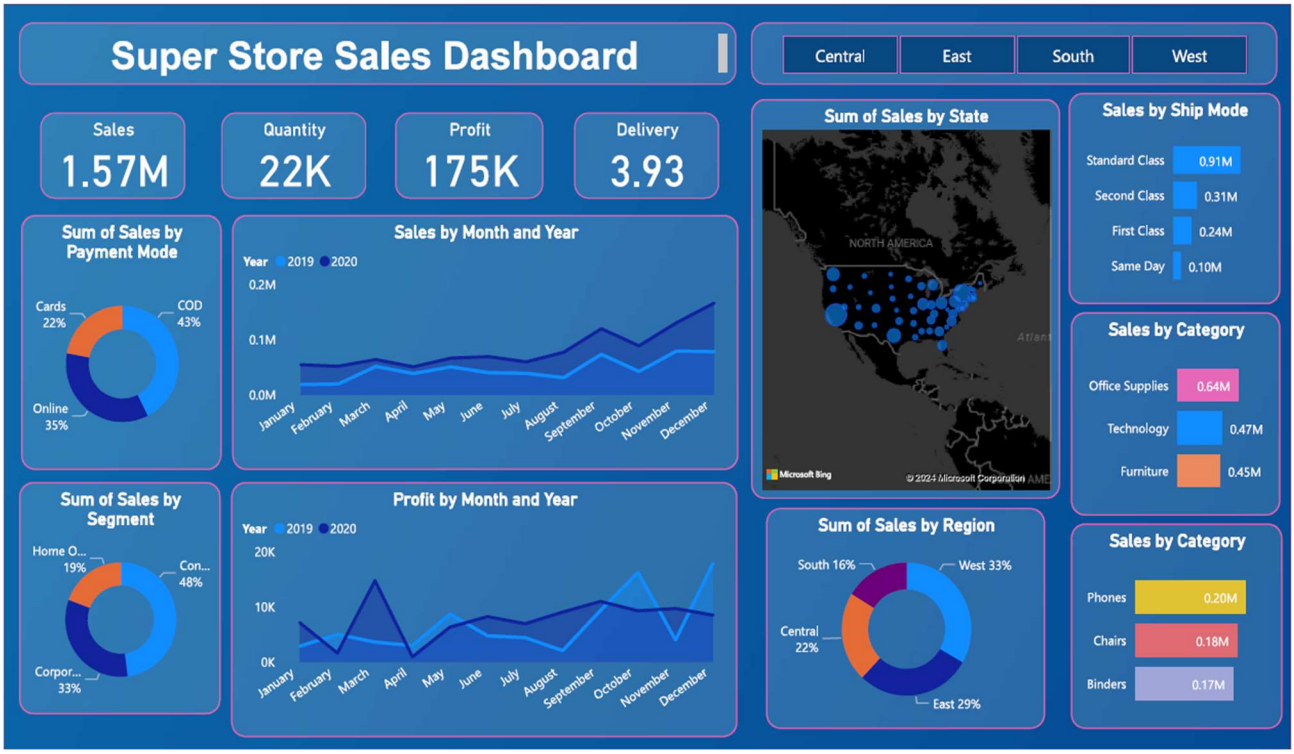
1. **Exploratory Data Analysis (EDA)**:
    - Understand sales trends, KPIs, and anomalies using visualizations.
2. **Interactive Dashboard**:
    - Create visuals for sales trends, forecasts, and customer insights.

---

8. Dashboard Design

Key Dashboard Elements:

1. **Sales Overview**:
    - Visualizations of total sales by region, category, and time.
    - Monthly and quarterly performance trends.
2. **Customer Insights**:
    - Segmentation by Corporate, Consumer, and Home Office.
    - Top customers based on revenue.
3. **Time Series Forecasting**:
    - Predicted sales for the next 15 days.
4. **Logistics Analysis**:
    - Delivery timelines and return rates by region.

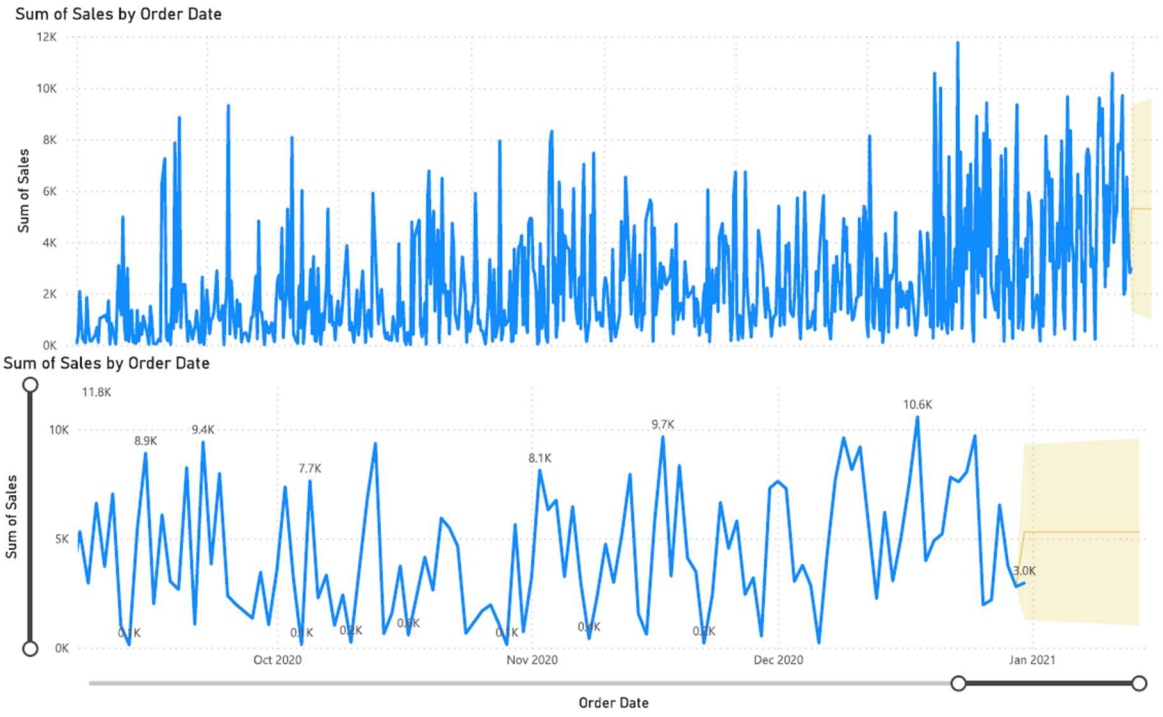

Super Store Sales Dashboard
Central
East
South
West
Sales
1.57M
Quantity
22K
Profit
175K
Delivery
3.93
Sum of Sales by State
NORTH AMERICA
Microsoft Bing
© 2024 Microsoft Corporation
Sales by Ship Mode
Standard Class 0.91M
Second Class 0.31M
First Class 0.24M
Same Day 0.10M
Sum of Sales by Payment Mode
Cards 22%
COD 43%
Online 35%
Sales by Month and Year
Year 2019 2020
0.2M
0.1M
0.0M
January February March April May June July August September October November December
Sales by Category
Office Supplies 0.64M
Technology 0.47M
Furniture 0.45M
Sum of Sales by Segment
Home O... 19%
Con... 48%
Corpor... 33%
Profit by Month and Year
Year 2019 2020
20K
10K
0K
January February March April May June July August September October November December
Sum of Sales by Region
South 16%
West 33%
Central 22%
East 29%
Sales by Category
Phones 0.20M
Chairs 0.18M
Binders 0.17M

9. Sales Forecasting

Process:

1. Aggregate sales data by day/month.
2. Generate a 15-day forecast to highlight peaks and dips in sales.

Output:
A detailed chart showing predicted sales with confidence intervals, actionable insights for inventory management.



10. Insights and Recommendations

Key Insights:

- **Regional Performance**: Tier 2 cities show growing demand but lag in delivery efficiency.
- **Product Trends**: Furniture and Technology categories drive the highest revenue.
- **Customer Preferences**: Online payments dominate; cash transactions declining.

Recommendations:

1. Focus promotional efforts on Tier 2 cities.
2. Partner with logistics providers to improve delivery in underperforming areas.
3. Increase stock availability for high-demand products.

---

11. Storytelling

Narrative Flow:

1. **Introduction**: Highlight the challenges faced and objectives of the analysis.
2. **Insights**: Show trends in sales and coordination with supporting visualizations.
3. **Impact**: Demonstrate the effect of recommendations on business growth.

---

12. Learning and Conclusion

This project has enhanced my understanding of:

- Time series forecasting and its applications in business.
- Creating interactive dashboards that simplify decision-making.
- Translating raw data into actionable insights to solve real-world problems.

Outcome: A well-rounded analysis that drives efficiency, increases profitability, and aligns with business objectives.