



# Superstore Sales Analysis & Forecasting Report

- **Project Overview**

To perform exploratory data analysis (EDA) and sales forecasting on a global superstore dataset with the goal of uncovering business insights and predicting short-term sales trends.

---

- **Dataset Overview**

The dataset contains 9,800 retail transactions from a global superstore over a period of four years (2015–2018). Each entry captures detailed information about:

Dates: Order date, Ship date

Customer: ID, name, segment

Geography: Region, state, city, postal code

Products: Category, sub-category, product name

Sales Info: Sales value, quantity, discount, and profit

Data Quality: Minor nulls in postal codes, handled. The dataset is otherwise clean and suitable for advanced analysis and forecasting.

---

- **Key Insights from EDA**

## 1. Year-over-Year Sales Growth

Year	Sales Change
2015 → 2016	↓ ~4%
2016 → 2017	↑ ~30%
2017 → 2018	↑ ~20%

### Interpretation:

Sales accelerated sharply after 2016, likely due to :

- Better marketing strategies
- High performing products
- An expanding and loyal customer base

---

## 2. Region-Wise Sales Performance

- **West:** ~\$700,000 (Highest)

- **East:** Strong second
- **Central:** Moderate
- **South:** ~\$390,000 (Lowest)

#### Interpretation:

- West and East are **key revenue hubs**.
- Central and South regions represent **growth opportunities** through better marketing and logistics optimization.

### 3. Category & Sub-Category Insights

#### By Category:

- **Technology:** > \$800,000
- **Furniture & Office Supplies:** Almost tied

#### Top Sub-Categories:

- **Phones & Chairs:** > \$300,000 each
- **Storage, Tables, Binders:** Solid contributors
- **Copiers & Accessories:** Meaningful
- **Art, Envelopes, Labels, Fasteners:** Low-performing

#### Interpretation:

- **Technology** is the most in-demand and high-margin category.
- Focus inventory & promotions on high-performing sub-categories.
- Evaluate the profitability of niche/low-sales items for possible phase-out.

### 4. Customer Segment Analysis

- **Consumers:** ~\$1.1M (Primary revenue driver)
- **Corporate:** ~\$670K
- **Home Office:** ~\$430K

#### Interpretation:

- Focus on **consumer loyalty programs** and **corporate partnerships** to boost overall sales.

### Time Series Forecast (Prophet Model)

Using Facebook's Prophet model, a **7-day forecast** was generated based on daily sales data.

Forecast Summary	Value (Approx.)
Forecast Period	Jan 1–7, 2019
Avg Predicted Sales	\$2,008/day
Trend	Stable, moderate variability

## Interpretation:

- Sales are projected to remain **stable**.
  - Helps guide **moderate restocking**, **resource planning**, and **staffing** for short-term operations.
- 

## Recommendations

- **Double down on Tech & Phone categories** — high ROI.
  - **Target South/Central regions** with localized promotions.
  - **Segment-based campaigns** for Home Office customers.
  - Use **forecasting outputs** to optimize stock and supply chain for the coming week.
- 

## Tools Used

- **Excel**: EDA, KPI dashboard, data cleaning
- **Python (Pandas, Seaborn, Prophet)**: Visual analytics & forecasting
- **Matplotlib**: Visualization
- **Jupyter Notebook**: Workflow