

DATA SCIENCE

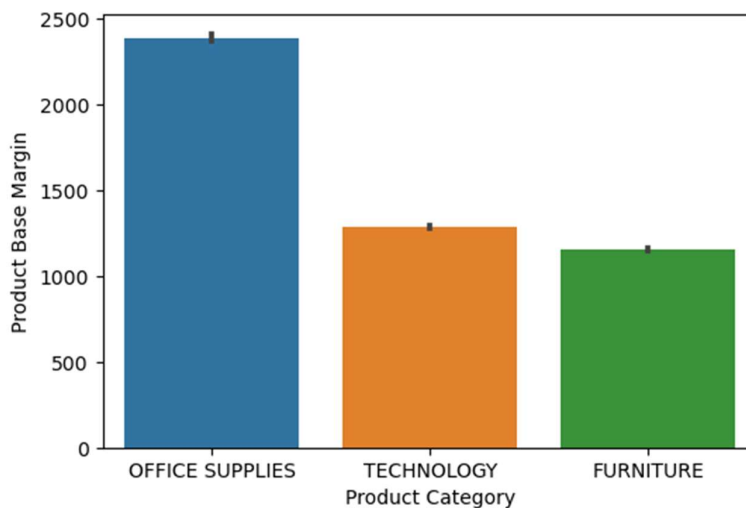
E-Commerce Sales Data Analysis

1.Introduction

This report provides a comprehensive analysis of categorical data from the sales dataset, focusing on key variables such as product categories, customer segments, and shipping modes. Along with the exploratory data analysis (EDA), Also present actionable insights derived from sales forecasting using time series analysis. The goal is to support the business team in making data-driven decisions related to marketing, inventory management, and strategic planning.

2. Exploratory Data Analysis (EDA)

➤ Product Base Margin



- **High Margin:** Categories like **Technology** and **Office Supplies** have higher margins, making them profitable.

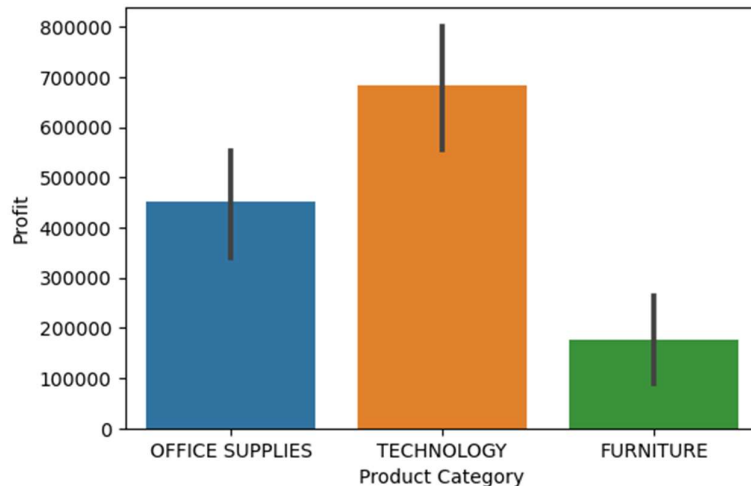
Action: Focus on promoting and expanding these categories to maximize profit.

- **Low Margin:** **Furniture** has a lower margin, likely due to higher manufacturing and shipping costs.

Action: Improve cost efficiency or adjust pricing strategies to increase margins.

By focusing on high-margin products and optimizing costs in low-margin areas, the business can enhance overall profitability.

➤ Profit by Product Category

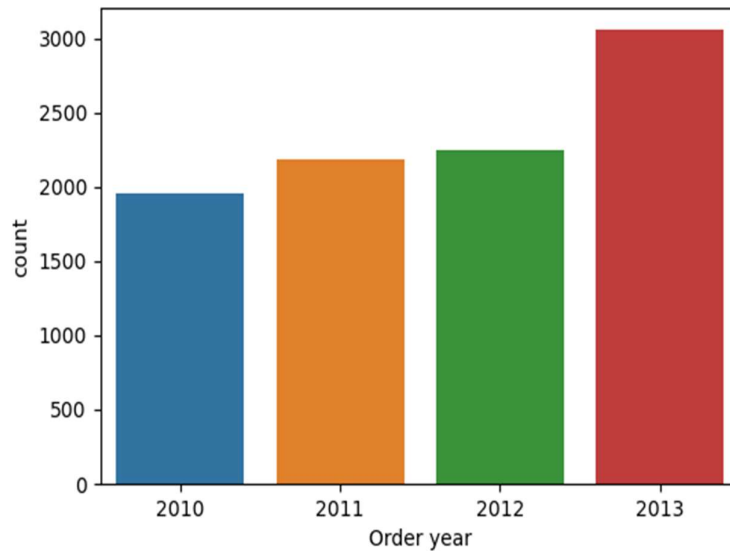


- **Insight:** The bar plot shows which product categories are contributing the most to the overall profit.
 - **Office Supplies** and **Technology** categories generate significant profits.
 - **Furniture** shows lower profit compared to the other categories.
- **Actionable Insight:** The business could focus on promoting products in the **Furniture** category to boost sales and profitability, or reconsider underperforming items. On the other hand, investing more resources into **Technology** products could maximize profit margins.

➤ Top 5 States by Sales Volume

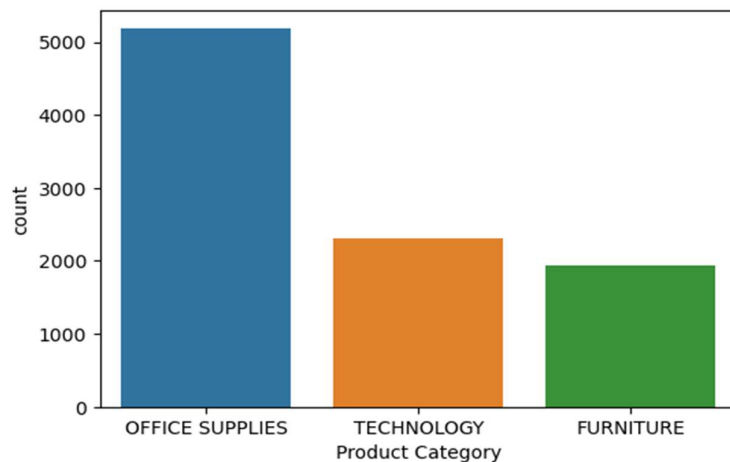
- **Insight:** The top states in terms of sales volume are **California, Texas, Illinois, New York, and Florida**.
 - These states are key contributors to the company's revenue.
- **Actionable Insight:** The company should focus on maintaining and expanding its presence in these states by introducing targeted marketing campaigns, improving customer service, or offering promotions in these regions.

➤ Order Count by Year



- **Insight:** There is a steady growth in the number of orders from 2010 to 2013, with a peak in 2013.
- **Actionable Insight:** The increase in order volume indicates growing demand. The business can prepare for higher sales in upcoming years by increasing inventory, expanding warehouse operations, and improving order fulfilment processes. Understanding the reasons behind this growth (e.g., better products, improved customer service) can help sustain it.

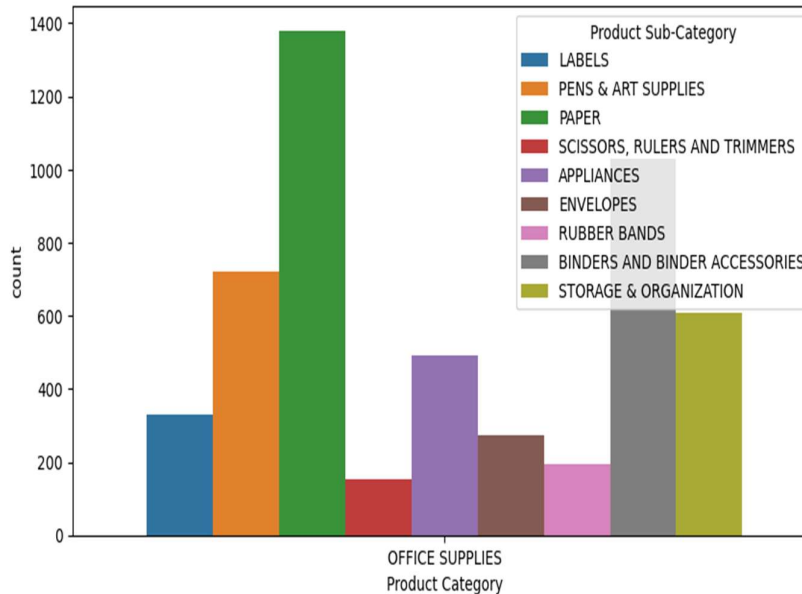
➤ Product Category Distribution



- **Insight:** The most ordered categories are **Office Supplies** and **Technology**, while **Furniture** has fewer orders.

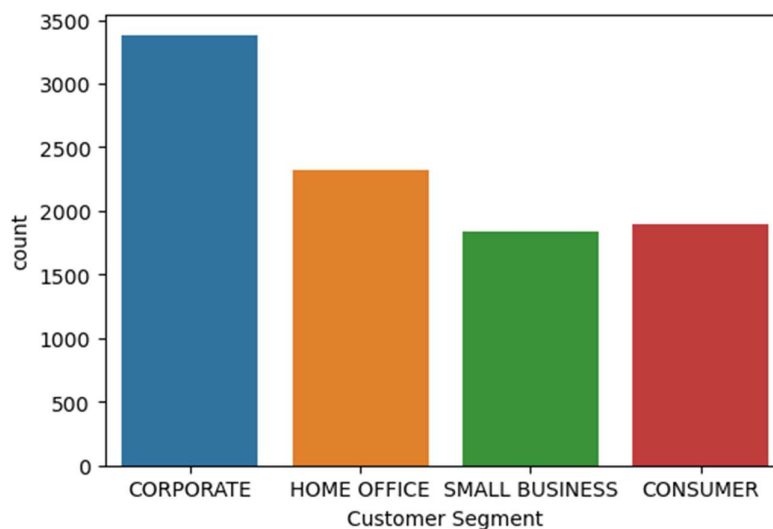
- **Actionable Insight:** To balance the product portfolio, the company could increase promotions for Furniture or bundle it with other popular categories. Analyzing customer preferences in these categories may also help refine the product offerings.

➤ Subcategory Distribution in Office Supplies



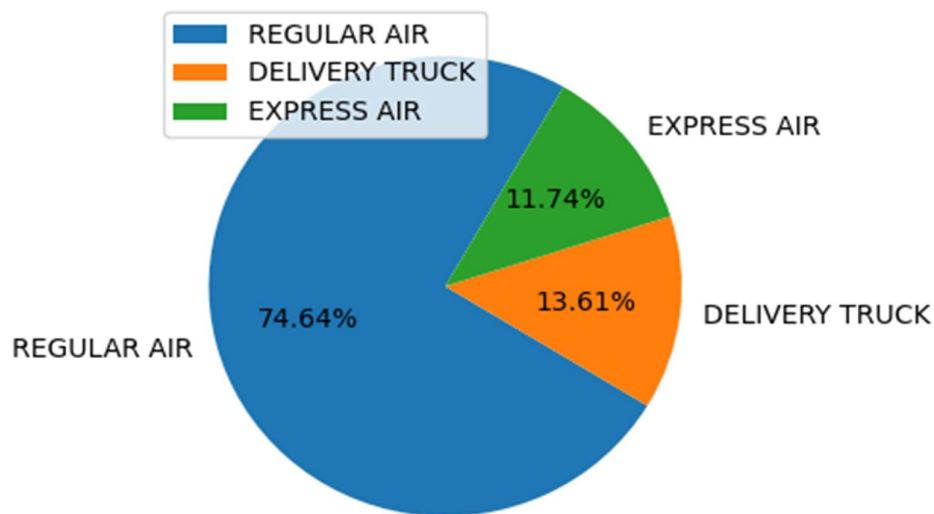
- **Insight:** Within **Office Supplies**, certain subcategories are more popular, such as **Binders and Binder Accessories**, **Pens & Art Supplies**, and **Paper**.
- **Actionable Insight:** The business can expand its offerings in these subcategories or introduce exclusive deals for these high-demand products. Low-demand subcategories could either be promoted or phased out.

➤ Customer Segment Distribution



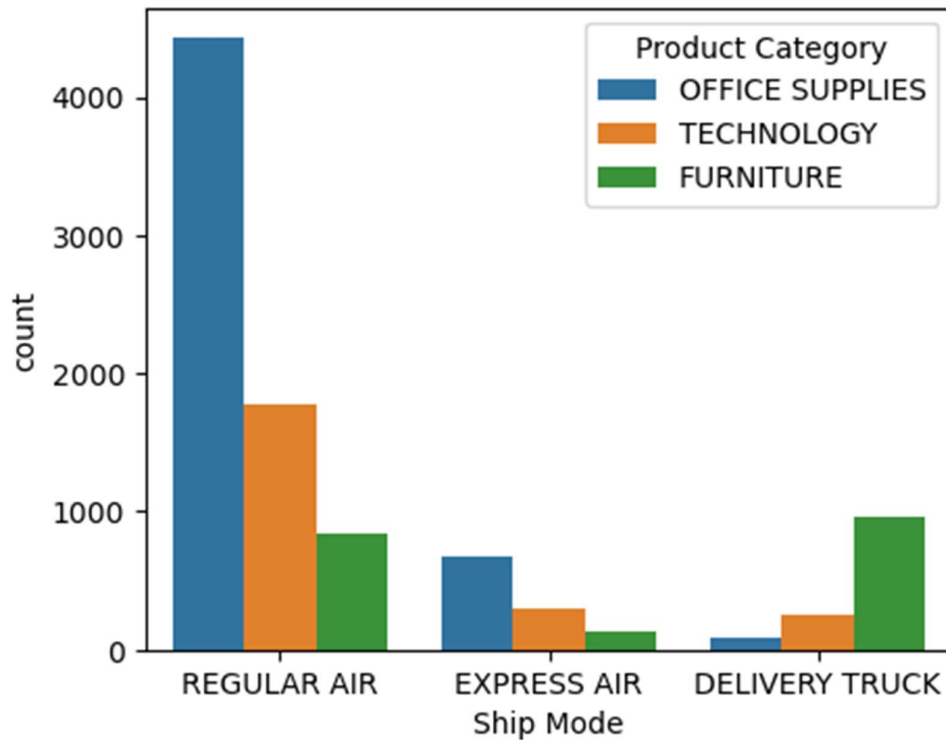
- **Insight:** The **Corporate** and **Home Office** segments represent the majority of the sales, with fewer orders from **Small Business** and **Consumer** segments.
- **Actionable Insight:** The company can introduce loyalty programs or customized promotions for **Small Business** and **Consumer** segments to drive more engagement. At the same time, nurturing relationships with **Corporate** clients is crucial, as they generate a significant portion of the revenue.

➤ **Shipping Mode Distribution**



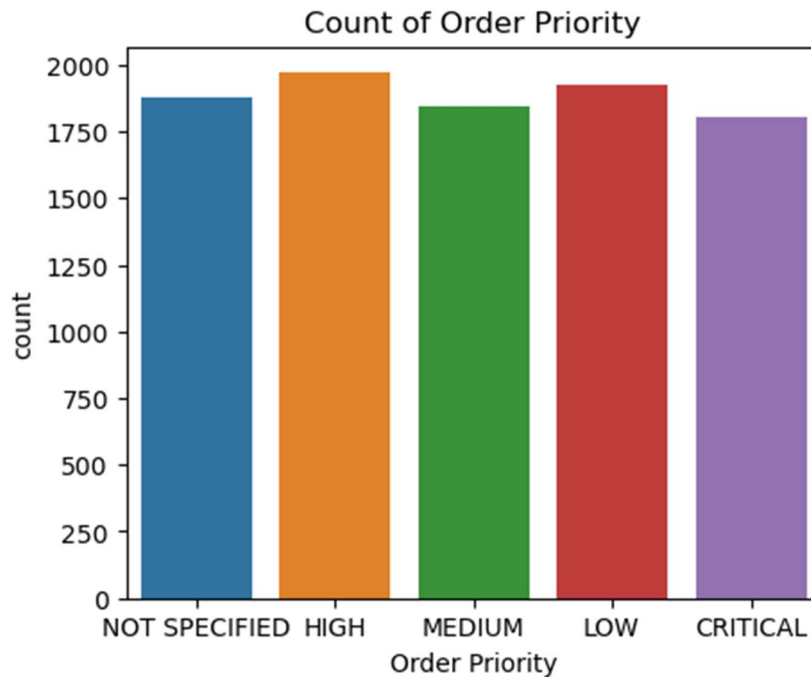
- **Insight:** The majority of the orders are shipped via **Regular Air** (74.64%), followed by **Delivery Truck** and **Express Air**.
- **Actionable Insight:** Since most orders are shipped via **Regular Air**, the company could evaluate the costs associated with this shipping mode. Exploring more cost-effective shipping options for certain products (without compromising delivery time) could lead to savings.

➤ **Shipping Mode by Product Category**



- **Insight:** The distribution of shipping modes across product categories shows that certain product categories (like **Office Supplies**) primarily use **Regular Air** shipping, while others may use **Express Air** for faster delivery.
- **Actionable Insight:** The company can optimize shipping strategies based on product categories. For high-margin products or time-sensitive products (like **Technology**), **Express Air** might be more appropriate, while for low-cost items (like **Office Supplies**), cost-saving methods like **Delivery Truck** can be used.

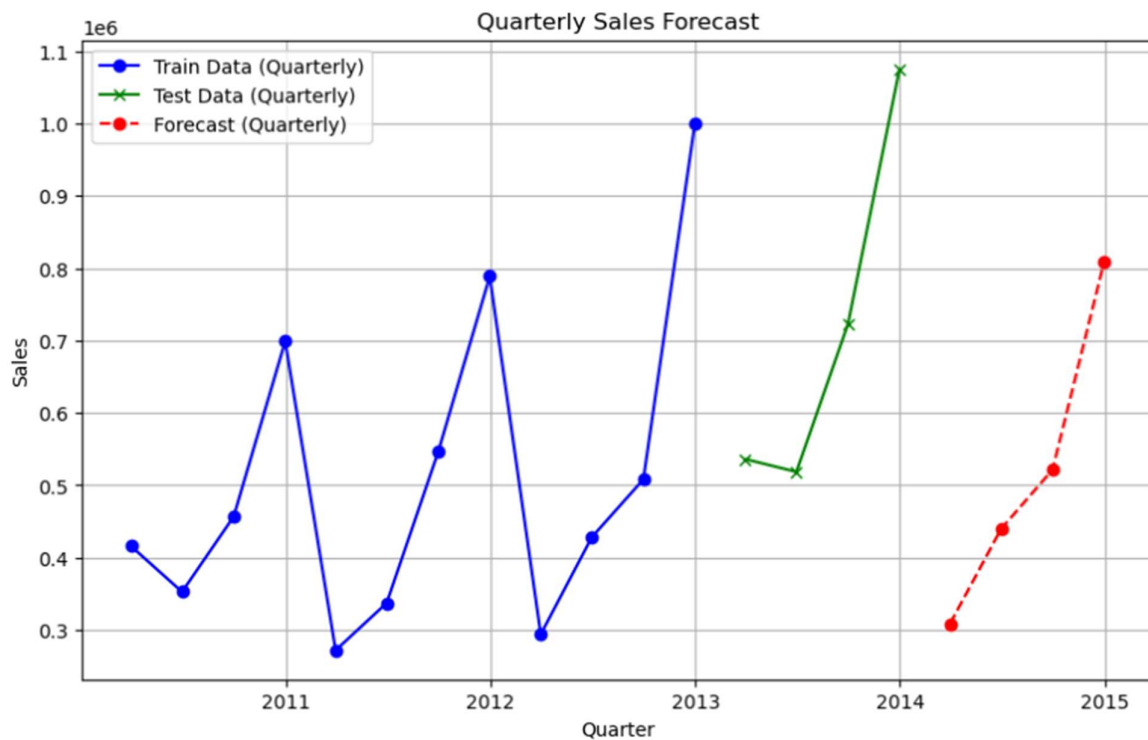
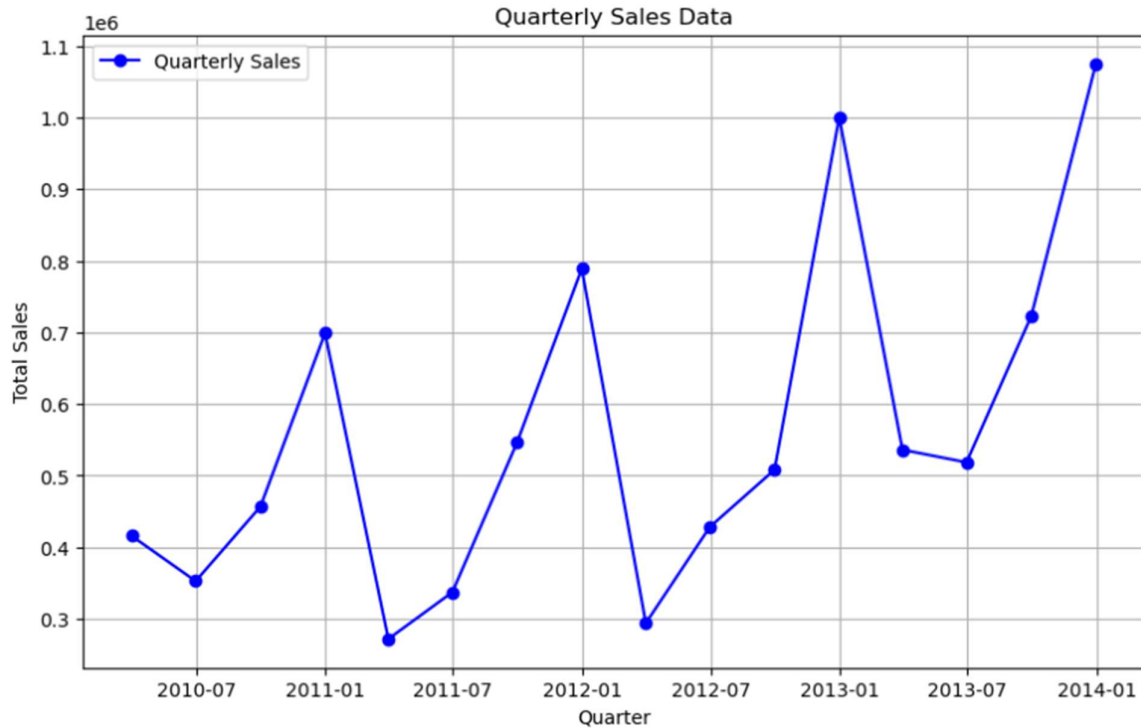
➤ Order Priority Distribution



- **Insight:** **High** and **Critical** priority orders make up a large portion of total orders, indicating the need for speed and reliability. A significant number of orders are categorized as **Not Specified**, potentially causing inefficiencies.
- **Actionable Insight:** Improve service levels for **High** and **Critical** orders by allocating more resources and prioritizing these orders in fulfilment. Address the **Not Specified** category by improving customer order processes to ensure clear priority selection.

3. Time Series Analysis (TSA)

➤ Quarterly Sales Analysis



Overview:

- The data is resampled on a quarterly basis, where the total sales for each quarter (Q1, Q2, Q3, Q4) are calculated.
- Quarterly sales are useful for identifying broader seasonal patterns and trends that repeat over the course of a year.

Time Period:

- The analysis includes data from **2010 to 2013**.
- For forecasting, the data from the first quarter of 2010 up to the fourth quarter of 2012 is used for training, while the sales from 2013 are used for testing.

Forecasting:

- A **4-quarter forecast** has been generated for the year **2014**, based on the data trends from 2010 to 2013.

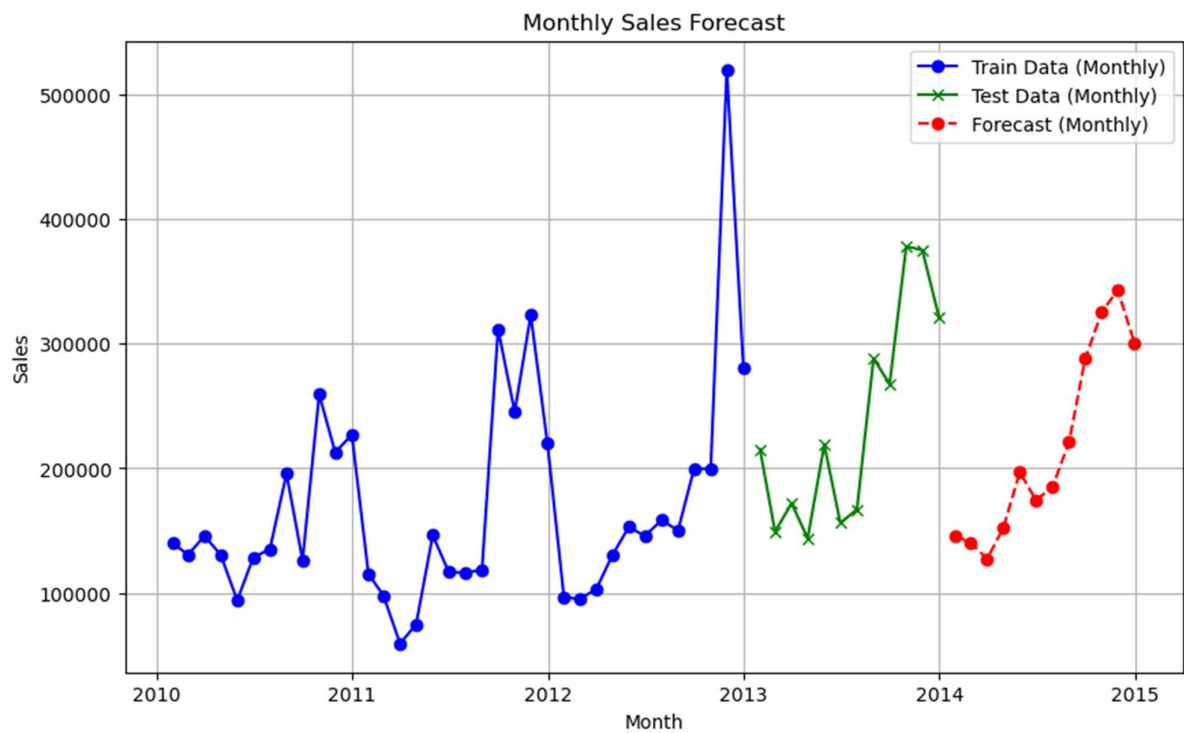
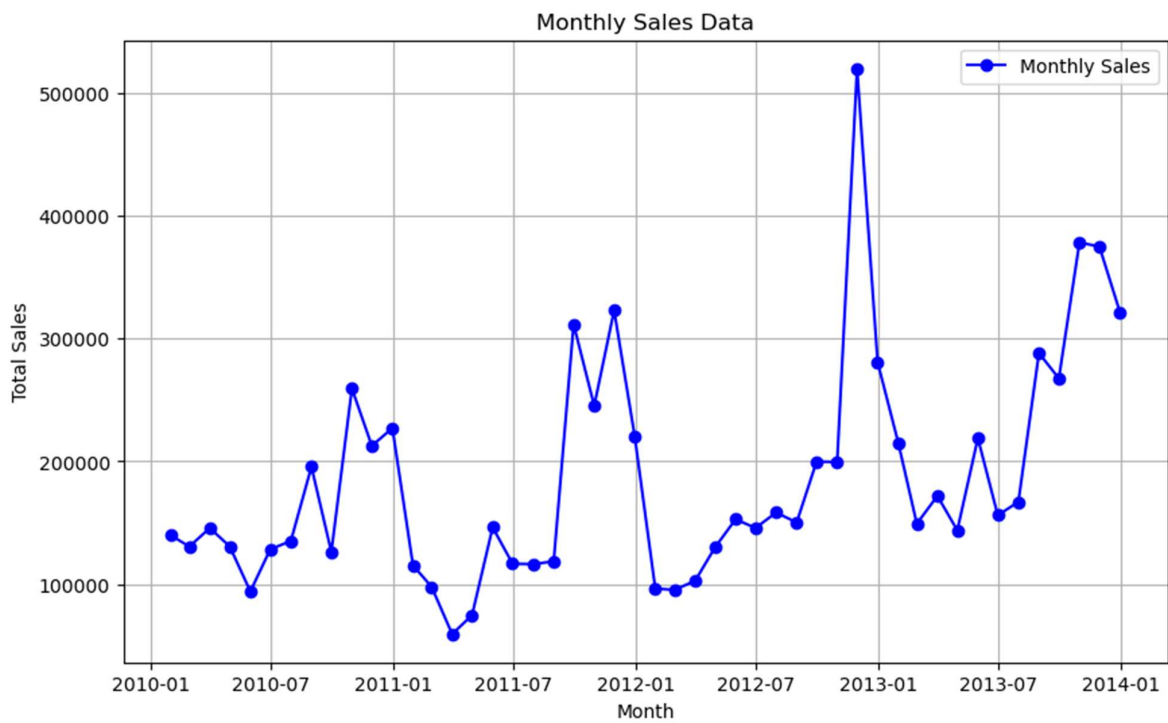
Graph Explanation:

- **Blue Line:** Represents actual sales from **2010 Q1 to 2012 Q4** (training data).
- **Green Line:** Represents actual sales in **2013** (test data).
- **Red Dashed Line:** Represents the forecasted sales for **2014**.

Business Insight:

- This quarterly forecast allows you to plan long-term strategies. For example, if sales peak in Q2 (April-June), the business can schedule major marketing campaigns or product launches during that time to maximize sales.

➤ Monthly Sales Analysis



Overview:

- This analysis captures monthly sales trends by resampling the data at monthly intervals.
- Monthly sales help identify more granular patterns, such as sales spikes driven by specific promotions, holiday seasons, or product launches.

Time Period:

- The analysis covers sales data from **January 2010 to December 2013**.
- The forecast is based on data from **January 2010 to December 2012** (training data) with testing on **2013** sales.

Forecasting:

- A **12-month forecast** has been generated for the year **2014**, allowing for monthly planning.

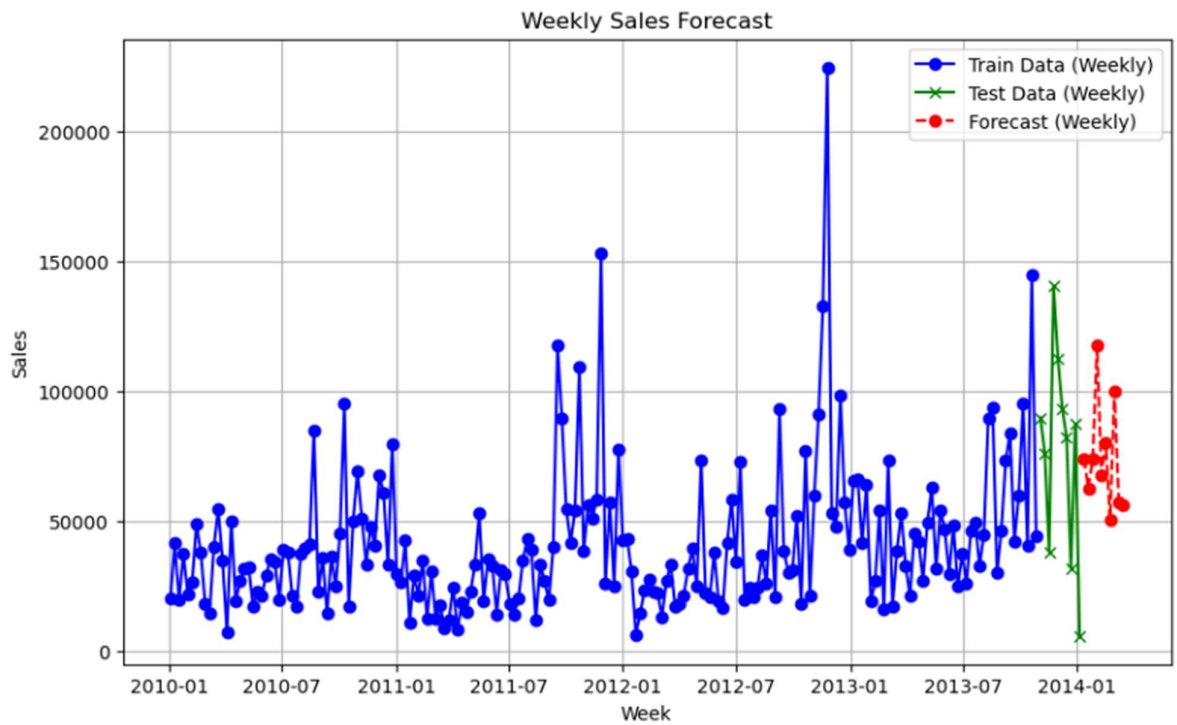
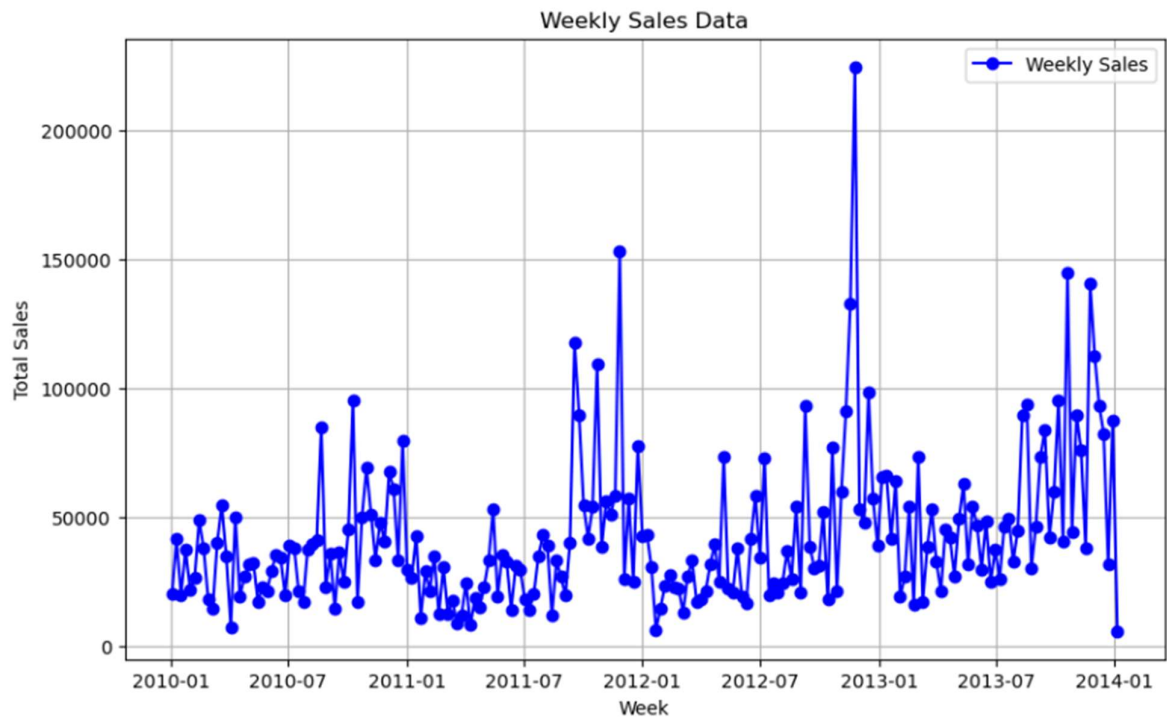
Graph Explanation:

- **Blue Line:** Represents actual sales from **January 2010 to December 2012** (training data).
- **Green Line:** Represents actual sales from **January 2013 to December 2013** (test data).
- **Red Dashed Line:** Represents forecasted sales for **2014**.

Business Insight:

- Monthly forecasts help plan operations, especially during key months like December (holiday season) or specific months that have consistently high sales. Adjusting stock levels and staffing based on monthly trends ensures smoother operations.

➤ Weekly Sales Analysis



Overview:

- Weekly sales trends offer even more granular insights, particularly useful for short-term planning.
- By resampling sales data weekly, we can see fluctuations that could be due to week-to-week marketing campaigns, short-term promotions, or external factors like public holidays.

Time Period:

- The analysis focuses on **2013 sales**, with data from **January 2013 to October 2013** used as training data and **November and December 2013** as the test period.

Forecasting:

- A **10-week forecast** has been generated for the first 10 weeks of **2014**, helping to plan for the immediate future.

Graph Explanation:

- **Blue Line:** Represents weekly sales from **January 2013 to October 2013** (training data).
- **Green Line:** Represents actual sales from **November to December 2013** (test data).
- **Red Dashed Line:** Represents forecasted sales for **January to March 2014**.

Business Insight:

- Weekly forecasts are critical for optimizing promotions, stock levels, and workforce management in the short term. For example, if sales are expected to rise in the coming weeks, the business can adjust by stocking up on inventory and increasing marketing efforts during those weeks.

4. Conclusion

The sales analysis reveals clear patterns in customer preferences, product demand, and regional sales performance. The data shows a concentration of sales in specific regions, like California and Texas, as well as high activity in product categories such as Office Supplies and Technology. Order fulfillment methods and customer segments also show distinct trends that can be leveraged to enhance operational efficiency and customer satisfaction. The time-series forecasting of sales offers a forward-looking view, allowing the business to prepare for future demand fluctuations.

5. Recommendations

1. Focus on High-Demand Regions:

- Prioritize marketing and promotional activities in top-performing regions such as California, Texas, and Illinois. This will maximize revenue opportunities by tapping into areas with proven sales success.

2. Enhance Inventory Management:

- Align inventory with sales forecasts to avoid overstock or stockouts, especially for high-demand categories like Office Supplies. Adjust stock levels dynamically based on regional demand and predicted sales trends.

3. Optimize Order Fulfillment:

- Streamline processes for high-priority and critical orders to improve customer satisfaction. Offering expedited delivery options for critical and high-value orders will enhance the customer experience.

4. Review Pricing Strategies:

- Reassess the pricing and profit margins of underperforming products. Consider adjusting prices, negotiating better supplier terms, or discontinuing low-margin products to improve overall profitability.

5. Strengthen Customer Segmentation Approach:

- Corporate customers show high activity. Developing loyalty programs, corporate contracts, and special pricing for the corporate segment could lock in recurring revenue and foster long-term relationships.

6. Improve Shipping Efficiency:

- Since Regular Air is the most utilized shipping method, evaluate its cost-efficiency. Offer faster, premium shipping options where applicable, and review the shipping performance in high-demand regions.

By implementing these recommendations, the business can better capitalize on existing trends, optimize operations, and enhance profitability while meeting customer expectations.