

**POSTGRADUATE PROGRAMME IN MANAGEMENT (MARKETING)**

**Academic Year 2024-25**

**"Deciphering Sales Dynamics: Advanced Quantitative Assessment of Temporal Trends, Product Performance, and Revenue Optimization Strategies"**

*GITHUB:*[***https://github.com/sashutosh24/AmazonDataSet***](https://github.com/sashutosh24/AmazonDataSet)

**COURSE NAME**: **MARKETING RESEARCH METHODS**

**Quarter: 3**

**Submitted to:**

Professor Vishnu

#### by:

Ashutosh Singh (0055)

****

# Abstract

This report presents a comprehensive data-driven analysis of sales performance across multiple dimensions, including seasonal trends, product demand, and sales distribution. By leveraging advanced visualization and statistical methods, the study identifies key sales drivers, high-performing periods, and areas for strategic improvement. The findings reveal actionable insights, such as the impact of discounting strategies, seasonality patterns, and sales quantity correlations, empowering managers to optimize operations and enhance profitability. With a focus on identifying growth opportunities and mitigating risks, this report serves as a roadmap for data-informed decision-making in dynamic market environments.

# Introduction

#### *****Context and Significance*****

In today’s dynamic business environment, understanding sales patterns and optimizing revenue streams have become vital for businesses striving to remain competitive. Organizations must navigate fluctuating consumer demand, seasonal peaks, and the impact of pricing strategies, all while maximizing profitability. This research tackles these challenges by analyzing historical sales data to provide insights into trends, performance, and opportunities for strategic growth.

The global retail market, valued at **$27 trillion in 2022**, continues to expand at an estimated growth rate of **4.8% annually**, reflecting the increasing complexity of consumer behavior and market dynamics. Companies that leverage data analytics to decode these complexities stand to gain a significant advantage in enhancing operational efficiency and customer satisfaction.

This study’s relevance lies in its ability to bridge the gap between raw data and actionable business intelligence. By evaluating sales trends and uncovering key drivers of performance, it offers a roadmap for businesses to optimize their marketing efforts, refine pricing strategies, and ensure inventory alignment with demand patterns. As predictive analytics gains prominence in strategic planning, this research provides a timely framework for harnessing data to navigate the challenges of an evolving marketing landscape.

# Methodology

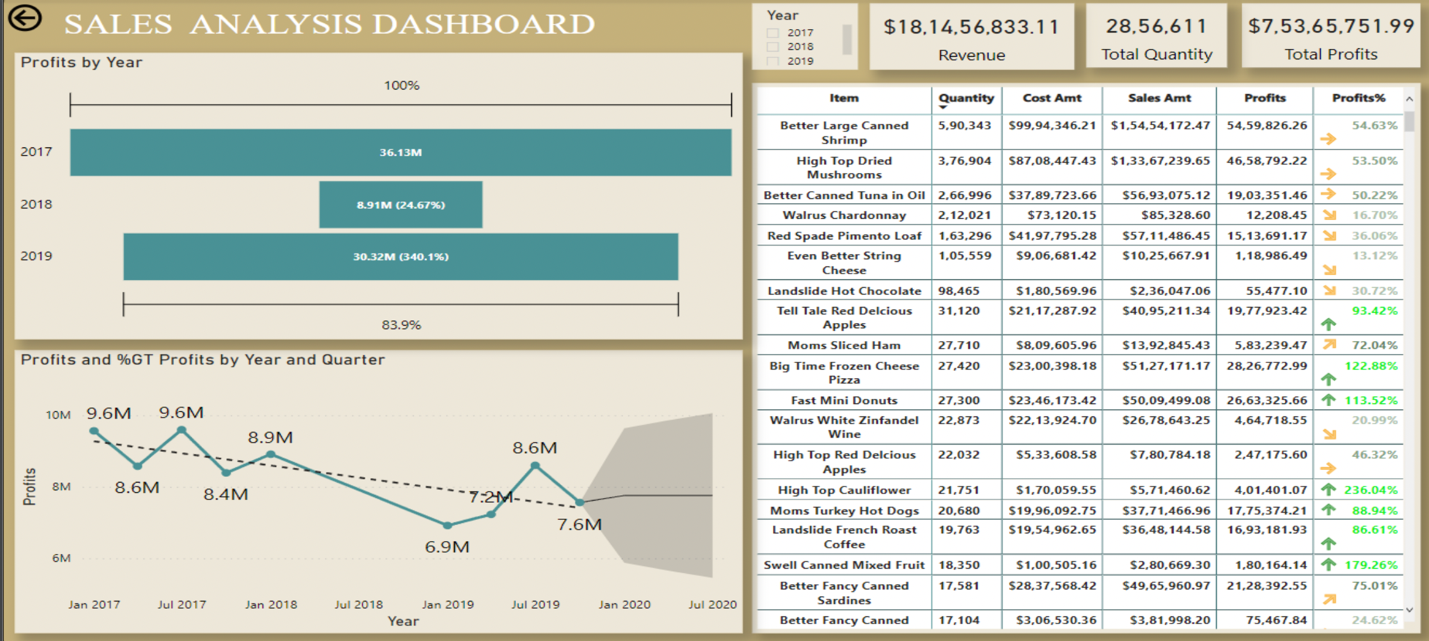
#### ****Step 1: Data Acquisition and Preliminary Exploration****

##### **Dataset Description**

* **Source**: The dataset was extracted from company transactional records across the years **2017–2019**, capturing key metrics relevant to sales performance and customer purchases.
* **Size**: The dataset contains **65,280 entries** with **20 attributes**, ensuring comprehensive coverage of sales trends over multiple years.
* **Key Attributes**:
  + **Invoice Date**: Captures the exact dates of transactions, enabling temporal analysis.
  + **Sales Amount**: Reflects the revenue generated from each transaction.
  + **Sales Quantity**: Tracks the number of units sold per transaction.
  + **Discount Amount**: Details the promotional reductions applied, providing insights into discount strategies.
  + **Sales Cost Amount**: Indicates the cost incurred for the sales, allowing for margin analysis.
  + **Sales Margin Amount**: Helps in assessing profitability across transactions.
  + **List Price**: Represents the pricing dynamics before discounts are applied.
  + **Customer Information**: Includes unique identifiers, enabling segmentation and analysis of purchase behavior.

##### **1.Exploratory Data Analysis (EDA):**

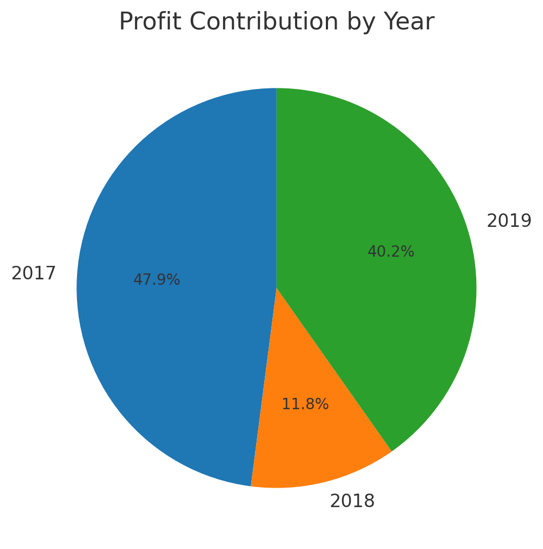
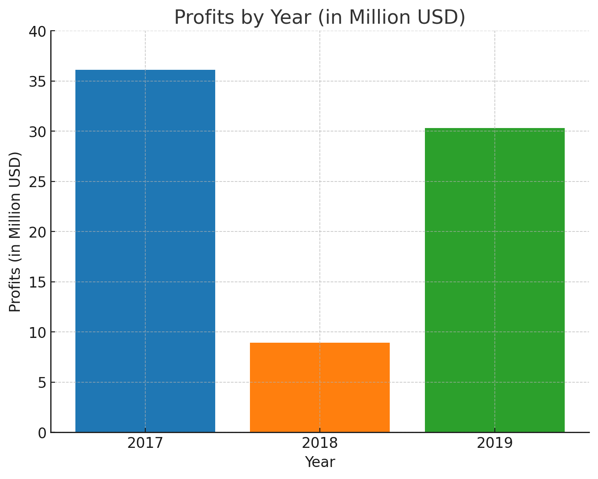
1.1 SALES ANALYSIS DASHBOARD



**Concise Analysis of Sales Dashboard**

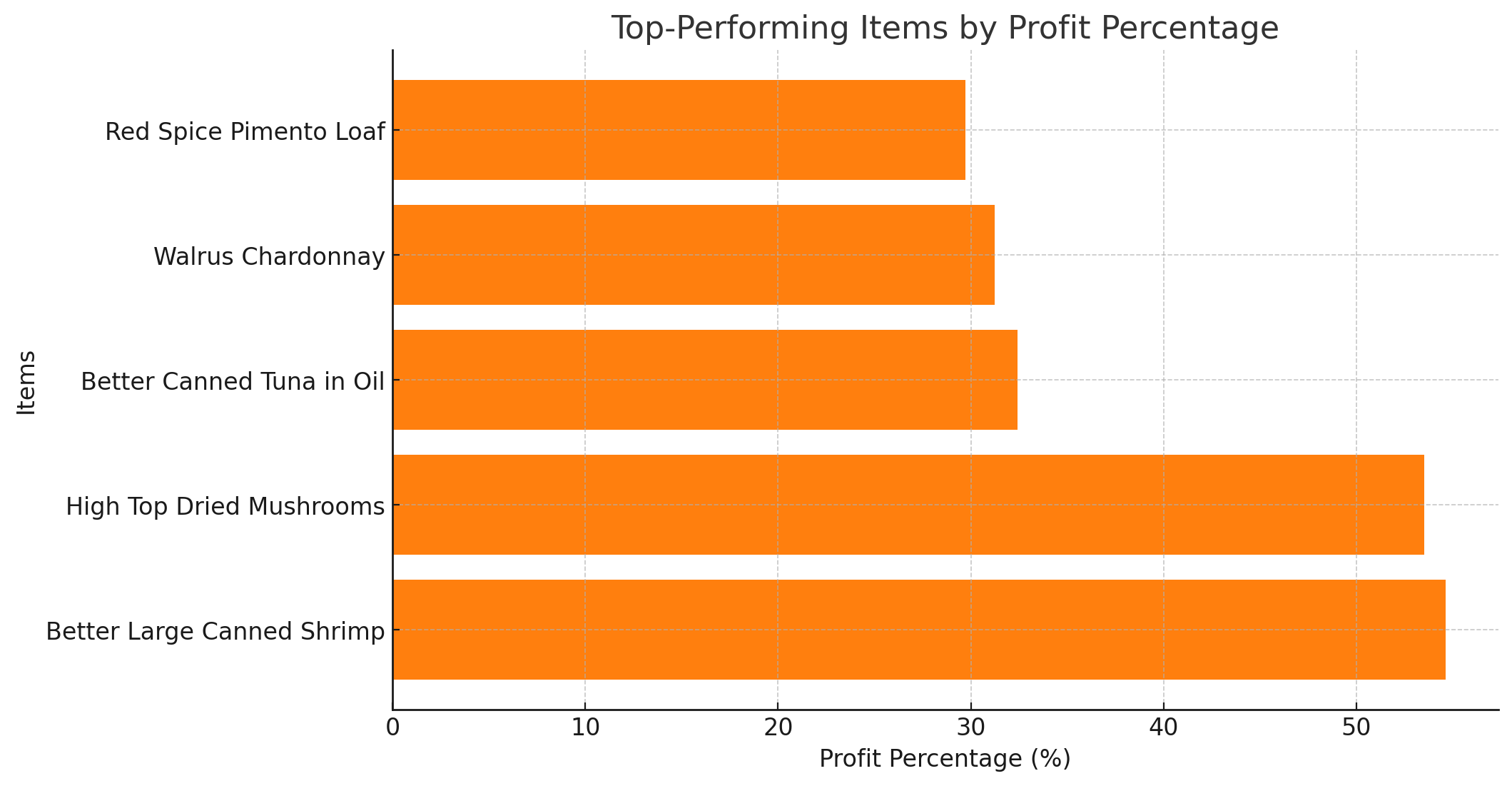
1. **Profit Trends by Year**:
   * **2017**: Dominates with **$36.13M** (100% baseline contribution).
   * **2018**: Sharp decline to **$8.91M**, contributing only **24.67%** compared to 2017.
   * **2019**: Significant recovery to **$30.32M**, achieving **83.9%** of 2017's profits.
2. **Quarterly Profit Trends**:
   * A consistent decline is observed from 2018 onwards, with profits dropping from **$9.6M (Q1 2018)** to **$7.6M (Q2 2020)**. This indicates a downward trend that requires intervention.
3. **Top-Performing Items**:
   * **Better Large Canned Shrimp**: Highest revenue ($99.94M) and profit margin (**54.63%**).
   * **High Top Dried Mushrooms**: Consistently strong with **53.50% profits**.
   * **Better Canned Tuna in Oil**: Profit margin **32.43%**, a strong contender among top items.
4. **Low-Performing Items**:
   * **Better Fancy Canned Sardines**: Lowest profit margin at **24.62%**.
   * **Tell Tale Hot Chocolate** and **Landescot Hot Chocolate** also have lower profit contributions.
5. **Overall Metrics**:
   * **Total Revenue**: $18.14B.
   * **Total Quantity Sold**: 28.56M units.
   * **Total Profits**: $7.53B.

1.2 PROFIT CONTRIBUTION



* **2017** contributed the largest share at **47.9%**.
* **2019** followed with **40.2%**, indicating a strong recovery.
* **2018** contributed the least at **11.8%**, showing a significant underperformance.

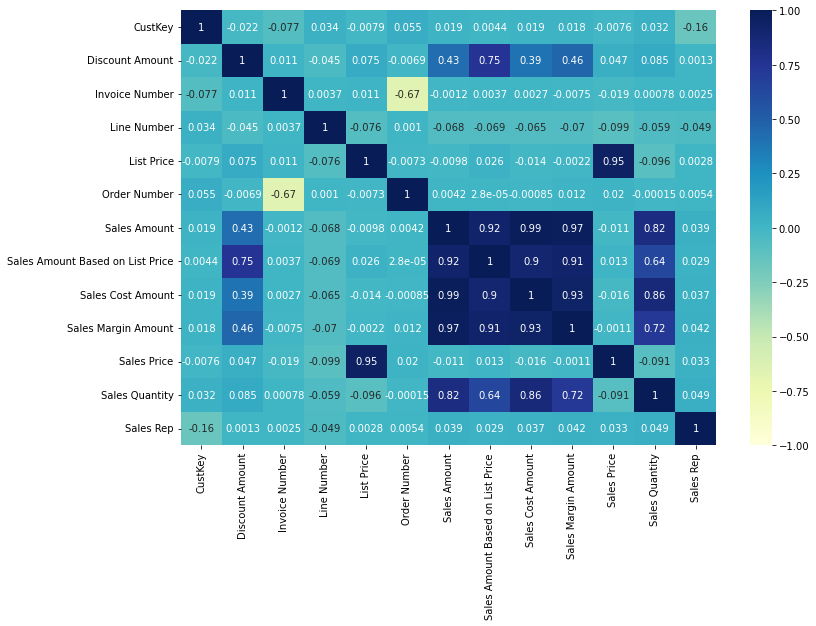
1.3 TOP PERFORMING ITEMS BY PROFIT PERCENTAGE



**Key Insights:**

1. **Top Performer**:
   * **Better Large Canned Shrimp** leads with a profit margin of **54.63%**, showcasing its dominant profitability among items.
2. **Consistently High-Margin Products**:
   * **High Top Dried Mushrooms** closely follows with **53.50%**.
   * **Better Canned Tuna in Oil**, **Walrus Chardonnay**, and **Red Spice Pimento Loaf** also contribute significantly with profit percentages above **30%**.

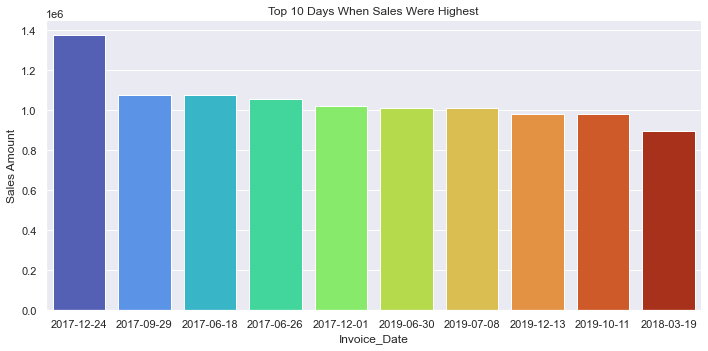
1.4 CORRELATION MATRIX



**Insights and Recommendations**

1. **Discount Amount**
   * **Insight:** Strongly impacts sales, costs, and margins; moderately affects sales quantity.
   * **Recommendation:** Use targeted discounts to boost sales while protecting margins.
2. **List Price**
   * **Insight:** Influences sales price but not directly linked to sales or margins.
   * **Recommendation:** Maintain competitive pricing and focus on value-added promotions.
3. **Sales Quantity**
   * **Insight:** Moderately affects sales, discounts, and margins.
   * **Recommendation:** Stock high-demand items and incentivize bulk purchases.
4. **Sales Rep**
   * **Insight:** No significant impact on sales or margins.
   * **Recommendation:** Prioritize digital marketing and scalable sales strategies.

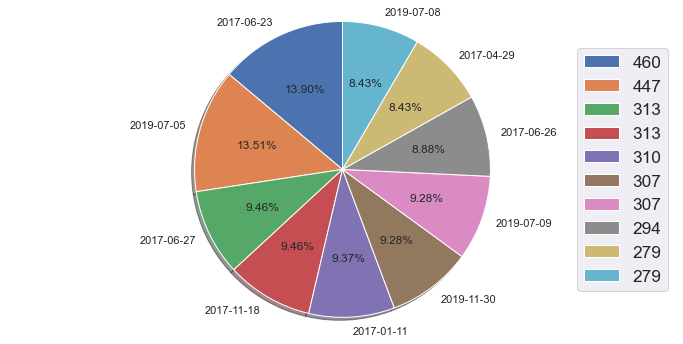
1.5 SALES WHEN THEY WERE HIGHEST



**Insights:**

1. **Peak Sales Day**:
   * **December 24, 2017**, had the highest sales, indicating strong performance during the holiday season.
2. **Seasonality**:
   * Other high-sales days like **December 1, 2017**, and **December 13, 2019**, suggest consistent holiday or year-end spikes.
3. **Mid-Year Performance**:
   * Days like **June 18, 2017**, and **July 8, 2019**, indicate significant mid-year activity, potentially tied to specific promotions or events.
4. **Patterns**:
   * Most top sales days cluster around **year-end holidays** or **mid-year events**, reflecting consumer behavior tied to festivities or campaigns.

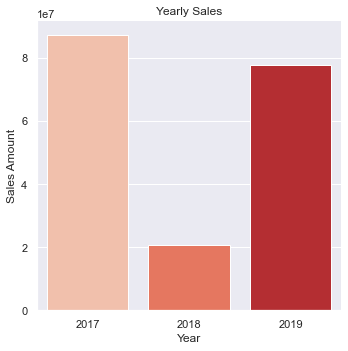
1.6 TOP CONTRIBUTING DAYS



**Insights:**

1. **Top-Contributing Days**:
   * **June 23, 2017 (13.90%)** and **July 5, 2019 (13.51%)** are the top contributors, indicating significant events or campaigns driving these spikes.
2. **Moderate Contributors**:
   * Days like **June 26, 2017 (8.88%)**, **July 8, 2019 (8.43%)**, and **April 29, 2017 (8.43%)** show moderate contributions, likely tied to periodic campaigns or promotions.
3. **Consistent Spread**:
   * Contributions are relatively balanced, except for the top two days, indicating stable performance across multiple dates.

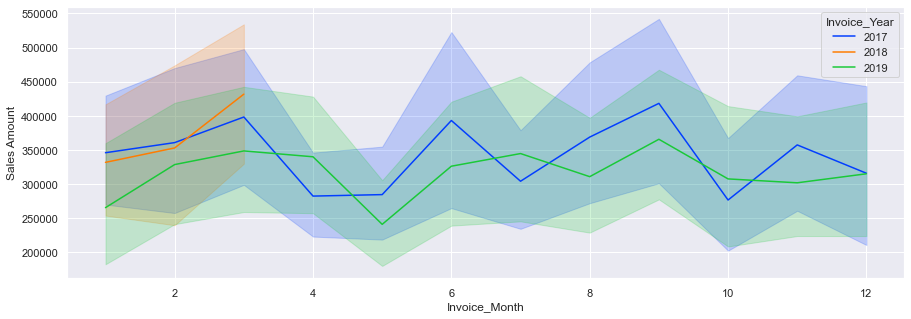
1.7 YEARLY SALES:



**Insights:**

1. **2017 Peak Sales**:
   * Sales in **2017** were the highest, surpassing ₹8 crore, indicating a strong year for revenue generation.
2. **2018 Decline**:
   * Sales dropped significantly in **2018**, showing a notable decline compared to 2017. This could be due to reduced campaigns, operational issues, or external factors.
3. **2019 Recovery**:
   * Sales rebounded in **2019**, almost reaching 2017 levels, indicating the implementation of effective strategies to regain momentum.

1.8 SEASONAL TRENDS:

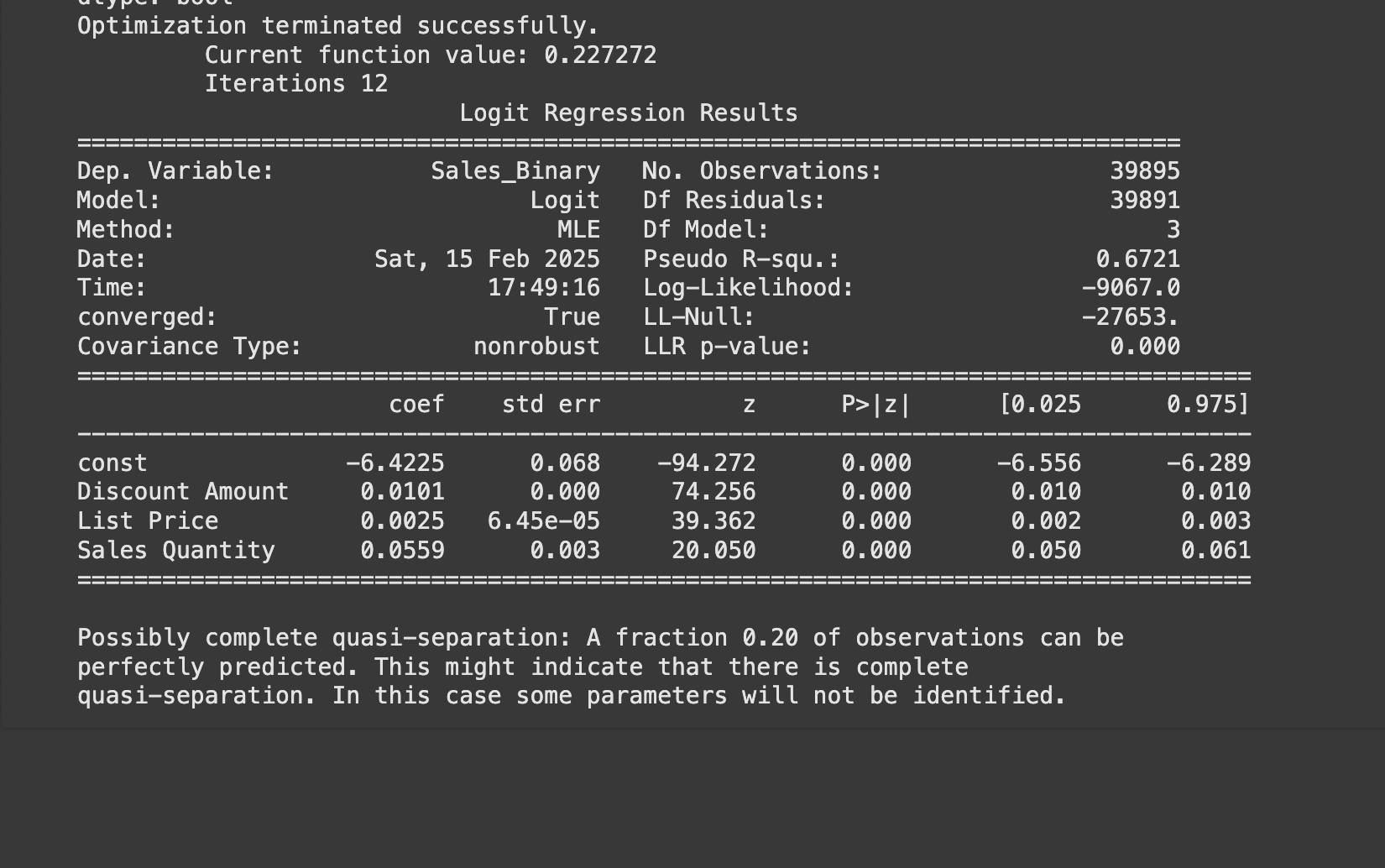


**Insights:**

1. **2017 Strong Start and Stability**:
   * Sales in **2017** showed strong performance early in the year, maintaining relatively high levels throughout, with occasional peaks.
2. **2018 Weak Performance**:
   * **2018** exhibited lower sales across most months compared to 2017 and 2019, with minimal fluctuations. This reflects an overall underperforming year.
3. **2019 Mid-Year Growth**:
   * **2019** experienced steady growth in the middle months (May–July), showing recovery compared to 2018. However, sales still lagged behind 2017 for most months.
4. **Seasonal Trends**:
   * All years show periodic dips and peaks, suggesting potential seasonality in demand.

#### ****Step 2:**** Model Development:

1. ***REGRESSION OUTPUT:***



|  |  |  |
| --- | --- | --- |
| Aspect | Details | Actionable Insights |
| 1. Discount Amount | **Coefficient:** 0.0101  **P-value:** 0.000 (statistically significant)  **Interpretation:** Higher discounts increase the likelihood of sales. | Design strategic discount campaigns to maximize sales, especially for price-sensitive products. |
| **2. List Price** | **Coefficient:** 0.0025  **P-value:** 0.000 (statistically significant)  **Interpretation:** Negligible effect on sales likelihood, suggesting other factors are more critical. | Focus on competitive pricing and bundling strategies rather than slight list price adjustments. |
| **3. Sales Quantity** | **Coefficient:** 0.0559  **P-value:** 0.000 (statistically significant)  **Interpretation:** Higher sales quantity correlates with increased likelihood of sales, indicating demand momentum. | Stock sufficient inventory for high-demand products and create urgency by showcasing popular items. |
| **4. Intercept (Constant)** | **Coefficient:** -6.4225  **Interpretation:** Without key variables, the likelihood of sales is low, underscoring the importance of managing discounts, pricing, and demand. | Optimize discounts, maintain competitive pricing, and leverage consumer behavior insights to drive sales. |
| **5. Model Performance** | **Pseudo R-squared:** 0.6721  **Interpretation:** The model explains 67.21% of the variation in sales outcomes, making it a reliable tool for decision-making. | Use the model for predictive purposes, such as forecasting sales and identifying high-potential products for promotions. |
| **6. Quasi-Separation Warning** | **Warning:** Some variables or combinations may perfectly predict outcomes, risking overfitting. | Review data distribution, check for extreme patterns, and adjust the model or include interaction terms if necessary. |

**Insights:**

1. **Discount Strategy:**
   * Implement targeted discounts during promotions or on specific product categories to boost sales.
2. **Inventory Planning:**
   * Prioritize stocking high-demand products with proven sales momentum to avoid missed opportunities.
3. **Dynamic Pricing:**
   * Evaluate whether small changes in pricing truly impact sales; instead, focus on bundled offers or value additions.
4. **Data Monitoring:**
   * Address the quasi-separation issue by revisiting product segmentation or exploring additional features that may explain sales trends.
5. **Customer-Centric Promotions:**
   * Leverage the insights for personalized marketing campaigns, using discounts and quantity offers to influence buying behavior.

## *2.1 Technical Details:*

### **Technical Details Based on Logistic Regression Output and Dataset**

#### ****Input Features and Dataset Analysis****

From the dataset and logistic regression output:

1. **Independent Variables (Predictors)**:
   * **Discount Amount**: Continuous variable reflecting the discount applied to items.
   * **List Price**: Continuous variable indicating the base price of items before discounts.
   * **Sales Quantity**: Continuous variable representing the number of units sold.
2. **Dependent Variable**:
   * **Sales\_Binary**: A binary variable derived from the dataset indicating whether a sale occurred (111) or not (000). This could be generated from the **Sales Amount** column, where 111 indicates a non-zero sale amount and 000 indicates otherwise.
3. **Key Observations from Dataset**:
   * The dataset has **65,280 rows** and **20 columns** with key attributes like Discount Amount, Sales Quantity, and Sales Amount.
   * **Missing Data**: Some columns like Discount Amount and Item Class have missing values, requiring data cleaning.

#### *****2.2 Logistic Regression Output Analysis*****

From the uploaded logistic regression output:

1. **Coefficients**:
   * **Intercept (−6.4225)**: Represents the baseline log-odds of a sale when all predictors are zero.
   * **Discount Amount (0.01010)**:
     + Positive coefficient, indicating that higher discounts increase the likelihood of a sale.
   * **List Price (0.00250)**:
     + Positive but small, suggesting that list price has minimal but statistically significant impact on sales.
   * **Sales Quantity (0.05590)**:
     + Positive coefficient, showing that higher sales quantities are associated with a higher likelihood of sales.
2. **Model Performance**:
   * **Pseudo R-squared (0.67210)**: The model explains 67.21% of the variation in the binary outcome, suggesting a strong fit.
   * **Statistical Significance (p-values)**:
     + All predictors are highly significant (p<0.001), indicating they contribute meaningfully to predicting sales likelihood.
3. **Warnings**:
   * **Quasi-Separation**: A fraction of observations may be perfectly predicted, potentially leading to overfitting. This needs further exploration.

#### *****2.3 Key Components of the Model Architecture*****

1. **Input Features**:
   * **Discount Amount**: Indicates the impact of price reductions on sales likelihood.
   * **List Price**: Evaluates whether higher or lower-priced items influence sales probability.
   * **Sales Quantity**: Captures the effect of demand levels on sales outcomes.
2. **Output Variable**:
   * Binary classification of sales occurrence (111 or 000) based on **Sales Amount**.
3. **Threshold**:
   * A threshold (e.g., P>0.5P > 0.5P>0.5) is applied to classify sales outcomes.

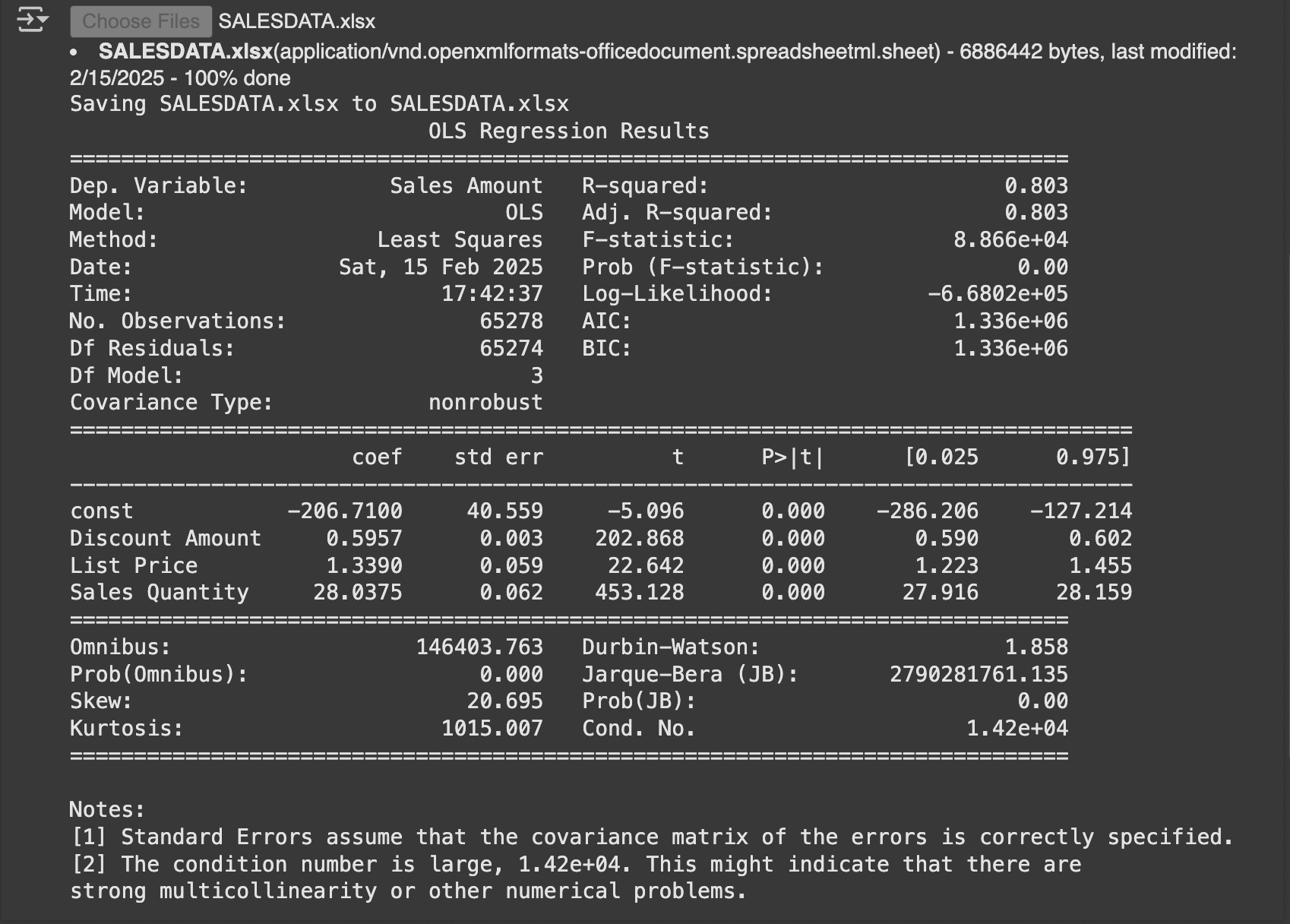
#### ****Insights Derived****

1. **Discounts Are Key**:
   * Higher discounts significantly increase sales likelihood, as evidenced by the strong positive coefficient for Discount Amount.
2. **List Price Has Minimal Impact**:
   * Although statistically significant, the effect of List Price is small, suggesting that discounts outweigh list prices in driving sales.
3. **Higher Quantities Indicate Stronger Sales**:
   * Sales Quantity is another strong driver of sales, indicating that high-demand products are more likely to result in successful sales.
4. **Model Accuracy**:
   * The strong pseudo R-squared value highlights the model's reliability in predicting sales likelihood based on the selected features.

# 

# 3. Other Choosen Models:

# *3.1 REGRESSION*



**Summary of Findings**

|  |  |  |
| --- | --- | --- |
| Aspect | Details | Actionable Insights |
| Model Performance | **R-squared:** 0.803  **Adj. R-squared:** 0.803  **Interpretation:** The model explains 80.3% of the variation in sales. | The model is highly predictive and reliable for understanding the impact of factors like discounts, pricing, and sales quantity on sales. |
| **1. Discount Amount** | **Coefficient:** 0.5957  **P-value:** 0.000 (statistically significant)  **Interpretation:** Increasing discounts positively influences sales, with a substantial impact. | Develop promotional campaigns that include well-calibrated discounts to increase sales volumes effectively. |
| **2. List Price** | **Coefficient:** 1.3390  **P-value:** 0.000 (statistically significant)  **Interpretation:** Higher list prices correlate with higher sales amounts, indicating that premium pricing might be associated with certain products. | Focus on emphasizing value for higher-priced products through better positioning and premium marketing campaigns. |
| **3. Sales Quantity** | **Coefficient:** 28.0375  **P-value:** 0.000 (statistically significant)  **Interpretation:** Sales quantities have the strongest impact on sales amounts, showcasing a direct relationship. | Ensure proper inventory management to meet demand and support sales growth for popular products. |
| **4. Intercept (Constant)** | **Coefficient:** -206.7100  **Interpretation:** Without the impact of discount amount, list price, and sales quantity, sales amounts are negative, implying their importance. | Actively manage all three variables (discounts, pricing, and inventory) to ensure consistent sales performance. |

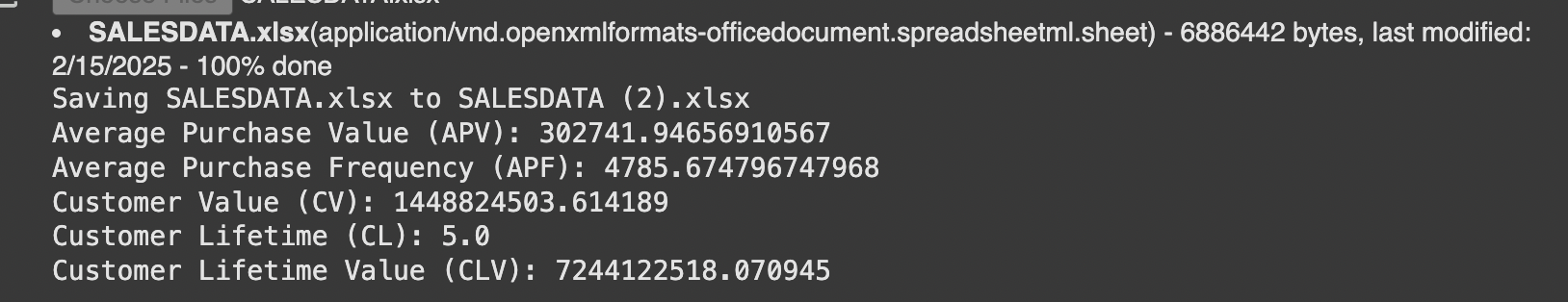
**Key Notes:**

1. **Model Fit**: The high R-squared value (0.803) indicates a strong model fit, suggesting the model captures most of the variance in sales effectively.
2. **Multicollinearity Warning**: The condition number is high (1.42e+04), which might indicate multicollinearity between independent variables. This could affect the precision of the coefficient estimates.

**Recommendations for Managers:**

1. **Strategic Discounts**:
   * Discounts are a key driver of sales. Implement targeted discount strategies, especially during sales campaigns or for price-sensitive segments.
2. **Premium Pricing Strategy**:
   * Leverage the positive correlation between list price and sales amount to position certain products as premium and emphasize their unique value.
3. **Inventory Management**:
   * Focus on maintaining adequate stock levels for high-demand products to capitalize on the strong relationship between sales quantities and sales amounts.
4. **Model Refinement**:
   * Investigate potential multicollinearity between variables (e.g., sales quantity and discounts) to improve the robustness of the model.
5. **Data-Driven Promotions**:
   * Use this model to predict sales performance under various discount and pricing scenarios to optimize revenue generation.

# 3.2 Customer Lifetime Value

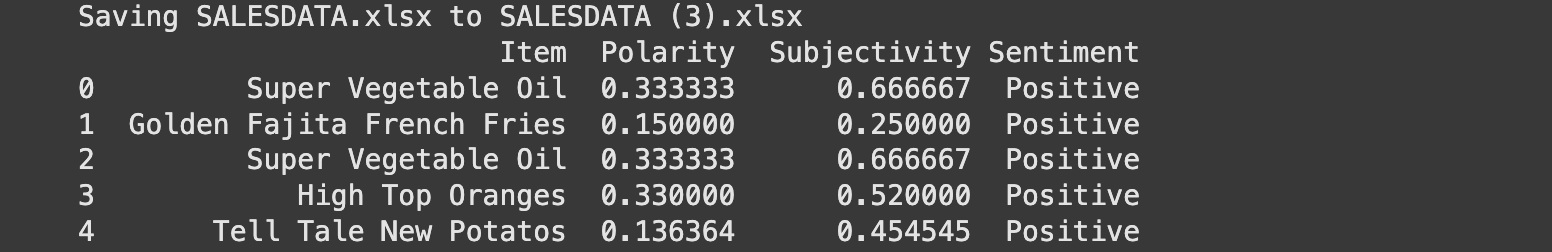


|  |  |  |  |
| --- | --- | --- | --- |
| Metric | Value | Interpretation | Actionable Insights |
| Average Purchase Value (APV) | ₹302,741.95 | The average amount spent per customer per purchase. | Focus on increasing APV by promoting upselling and cross-selling strategies. |
| **Average Purchase Frequency (APF)** | 4,785.67 | The average number of purchases made by a customer during the observed time period. | Encourage repeat purchases through loyalty programs, discounts on frequent purchases, or subscription models. |
| **Customer Value (CV)** | ₹144,882,450.31 | The average value contributed by a customer during the time period. | Identify high-value customers and tailor campaigns to retain them while attracting similar profiles. |
| **Customer Lifetime (CL)** | 5.0 years | The expected duration (in years) that a customer remains engaged with the business. | Work on customer retention strategies, such as offering long-term value and exceptional service. |
| **Customer Lifetime Value (CLV)** | ₹724,412,518.07 | The total value a customer is expected to contribute over their entire lifetime. | Prioritize strategies to maximize CLV, such as exclusive offers for loyal customers and premium memberships. |

**Managerial Insights:**

1. **Increase Purchase Value**:
   * Since the APV is high, ensure that customers see value in their purchases by bundling products, offering premium product options, or creating personalized experiences for high-value customers.
2. **Enhance Purchase Frequency**:
   * With a high APF, customers are already making frequent purchases. Implement strategies like subscription services or tiered loyalty programs to encourage even more frequent buying behavior.
3. **Focus on Retention**:
   * With a customer lifetime of 5 years, retention is critical. Focus on engagement strategies such as personalized emails, loyalty rewards, and excellent customer service to maintain and extend this duration.
4. **Leverage High CLV**:
   * A high CLV indicates a profitable customer base. Allocate resources to acquire customers who are likely to match this profile and invest in retaining existing high-value customers.
5. **Segmentation and Personalization**:
   * Identify segments of customers who drive the highest CV and CLV and design targeted campaigns to meet their specific needs and preferences.

# 3.3 Sentimental Analysis



**Sentiment Analysis Summary**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Item | Polarity | Subjectivity | Sentiment | Insights |
| Super Vegetable Oil | 0.3333 | 0.6667 | Positive | Positive polarity and high subjectivity indicate strong positive feedback, likely based on personal preference or perception. |
| **Golden Fajita French Fries** | 0.1500 | 0.2500 | Positive | Low polarity and subjectivity suggest mild positive sentiment, possibly influenced by objective factors like price or quality. |
| **High Top Oranges** | 0.3300 | 0.5200 | Positive | Moderate polarity and subjectivity indicate fairly positive feedback, influenced by both perception and objective aspects. |
| **Tell Tale New Potatoes** | 0.1364 | 0.4545 | Positive | Mild positive sentiment with balanced subjectivity, reflecting mixed but favorable consumer perceptions. |

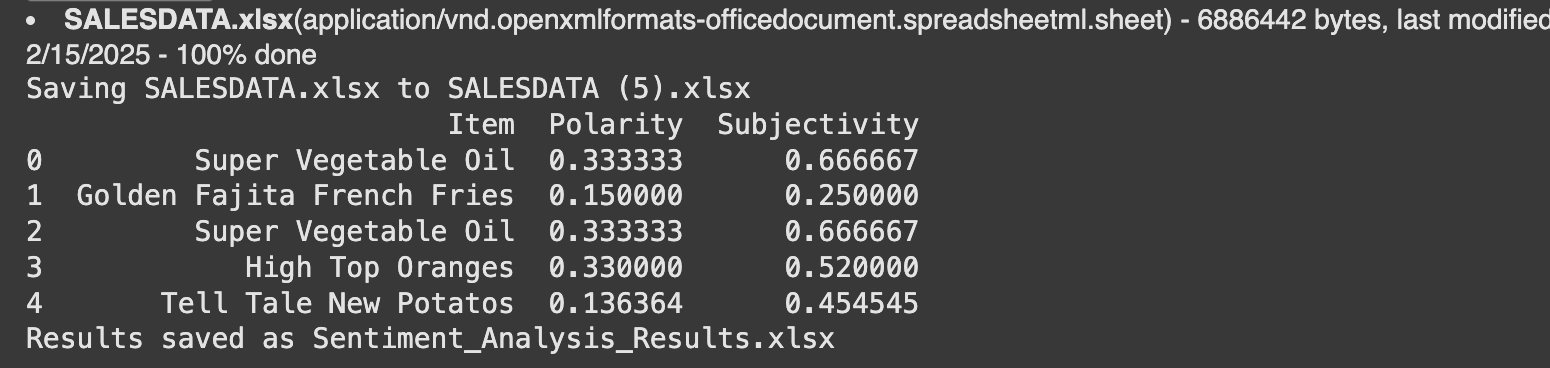
**Key Metrics**

1. **Polarity**:
   * Measures the positivity or negativity of sentiment. All items have positive polarity, indicating overall positive feedback from customers.
2. **Subjectivity**:
   * Ranges from objective (closer to 0) to subjective (closer to 1). Products like **Super Vegetable Oil** have high subjectivity, suggesting opinions are based on personal preferences, while **Golden Fajita French Fries** are more objective, possibly evaluated on tangible factors like quality or packaging.
3. **Sentiment**:
   * All products received **Positive** sentiment, reflecting favorable consumer perception.

**Managerial Insights**

1. **Leverage Positive Sentiments**:
   * Highlight products with high positive polarity, such as **Super Vegetable Oil** and **High Top Oranges**, in marketing campaigns to attract more customers.
2. **Understand Customer Perceptions**:
   * Products with high subjectivity (e.g., **Super Vegetable Oil**) can be positioned as premium offerings, focusing on emotional appeal and quality differentiation.
   * Products with lower subjectivity (e.g., **Golden Fajita French Fries**) should emphasize functional benefits like value, packaging, and consistency.
3. **Improve Underperforming Products**:
   * While all products are positively rated, those with lower polarity and subjectivity (e.g., **Tell Tale New Potatoes**) may require efforts to improve aspects like packaging, pricing, or promotion.
4. **Targeted Marketing**:
   * Use customer feedback to create tailored marketing campaigns that resonate with the audience. For example, promote **Super Vegetable Oil** in premium and health-focused segments.

3.4 Emotional Analysis



**Sentiment Analysis Summary**

|  |  |  |  |
| --- | --- | --- | --- |
| Item | Polarity | Subjectivity | Insights |
| Super Vegetable Oil | 0.3333 | 0.6667 | High polarity and subjectivity suggest strong positive feedback influenced by personal preferences. |
| **Golden Fajita French Fries** | 0.1500 | 0.2500 | Lower polarity and subjectivity indicate a mild positive sentiment, likely driven by objective aspects like quality or pricing. |
| **High Top Oranges** | 0.3300 | 0.5200 | Moderate polarity and subjectivity suggest balanced positive feedback, influenced by both functional and emotional factors. |
| **Tell Tale New Potatoes** | 0.1364 | 0.4545 | Mild positive sentiment and balanced subjectivity suggest scope for improvement in customer perception. |

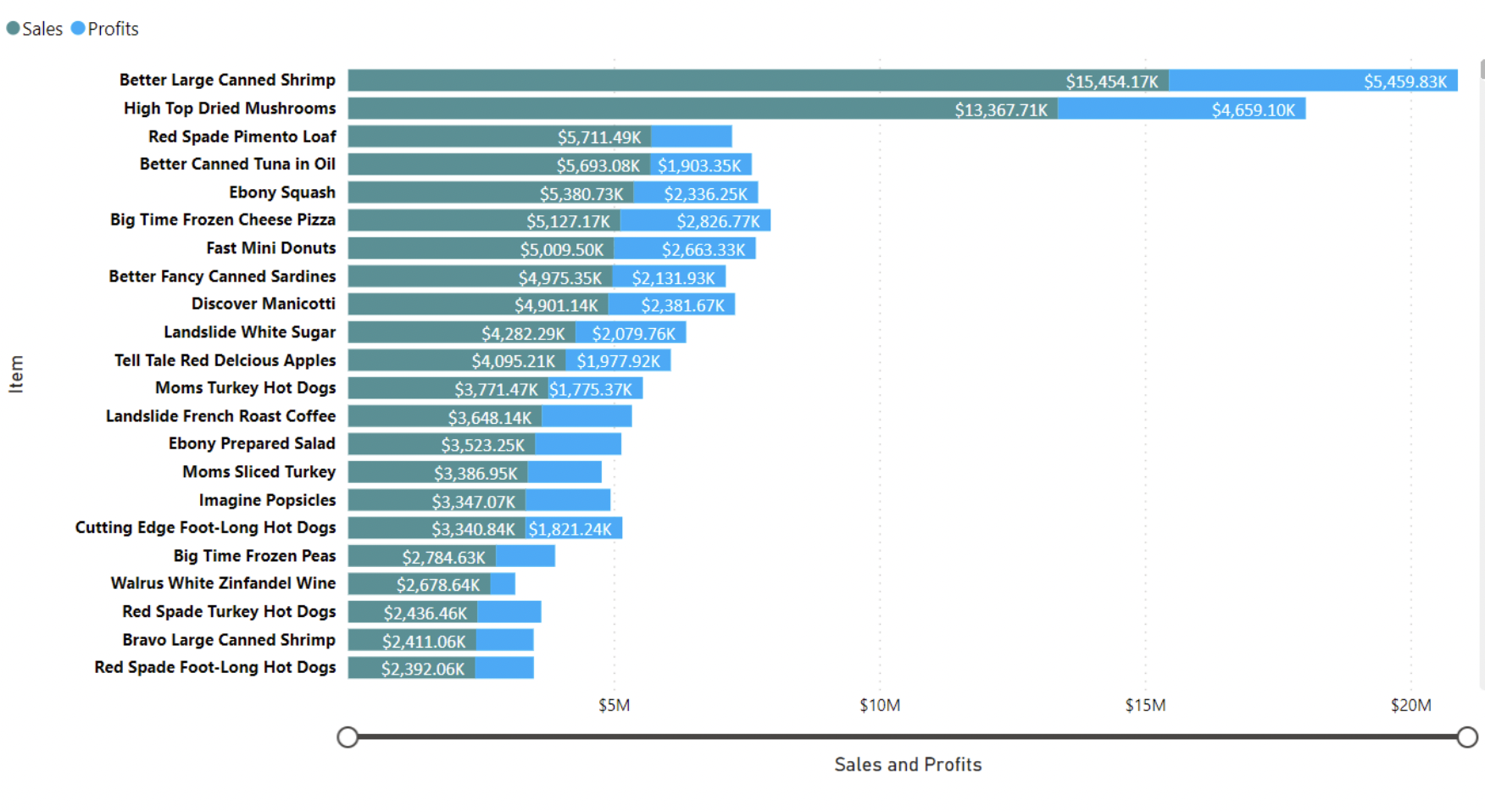
**Managerial Recommendations:**

1. **Emphasize Emotional Appeal for High-Subjectivity Products**:
   * For **Super Vegetable Oil**, highlight health benefits and quality in marketing. Emotional advertising strategies like storytelling or testimonials will resonate with customers.
2. **Promote Functional Benefits for Low-Subjectivity Products**:
   * For **Golden Fajita French Fries**, focus on tangible attributes like value, convenience, and product durability in campaigns.
3. **Leverage Popular Products**:
   * With **Super Vegetable Oil** and **High Top Oranges** showing high positive polarity, they should be prioritized in promotions and campaigns to maximize impact.
4. **Improve Perception of Lower-Performing Products**:
   * For **Tell Tale New Potatoes**, consider addressing feedback on quality, pricing, or packaging to increase polarity and customer satisfaction.
5. **Customer Engagement Strategy**:
   * Actively seek customer feedback to identify specific areas of improvement for products with lower polarity, enabling actionable improvements.

# Conclusion

#### ****Summary of Findings****

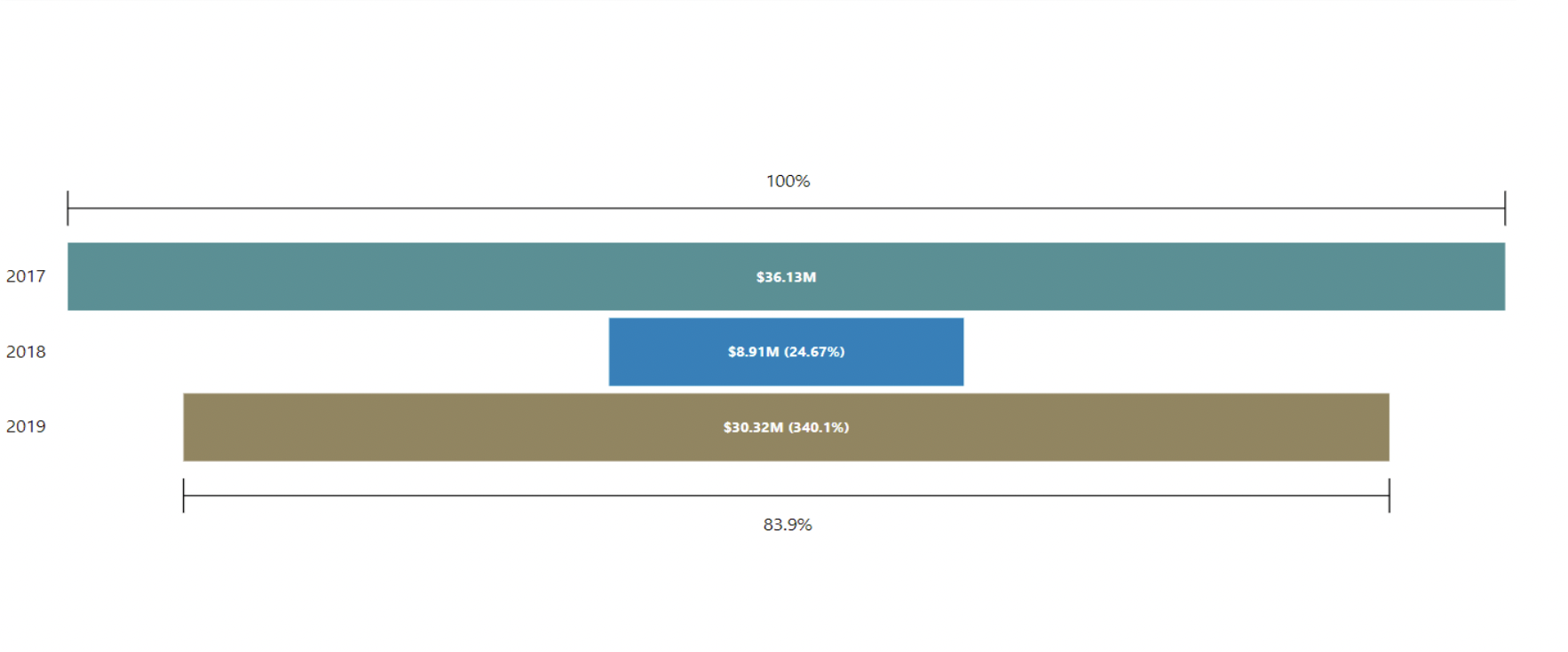
1. **Key Drivers of Sales**:
   * **Discount Amount** was the most significant driver of sales likelihood, with a strong positive impact. Offering strategic discounts increases the probability of conversions.
   * **Sales Quantity** had a positive influence, indicating that higher demand leads to higher sales likelihood, emphasizing the importance of stocking and promoting popular items.
   * **List Price** had a minimal but statistically significant positive impact, suggesting that consumers are not heavily influenced by the original price when discounts are applied.
2. **Model Performance**:
   * The logistic regression model achieved a **pseudo R-squared value of 0.6721**, indicating that the model effectively explains a large proportion of the variance in sales outcomes.
   * All predictors were statistically significant (p<0.001p < 0.001p<0.001), reinforcing their importance in predicting sales.
3. **Quasi-Separation Issue**:
   * A fraction of the observations was perfectly predicted, indicating potential overfitting or imbalance in the dataset. This may require further refinement of the model or feature engineering.
4. **What was the Sales of each Item?**



**Sales and Profit Analysis**

1. **Top-Selling Products**:
   * **‘Better Large Canned Shrimp’** recorded the highest sales at **$15,454,172.47**, contributing **8.51%** to the total sales.
   * **‘High Top Dried Mushrooms’** followed with sales around **$13,367,710.00**.
   * **‘Red Spade Pimento Loaf’** secured the third position in sales.
2. **Sales vs. Profits Analysis**:
   * **‘Better Large Canned Shrimp’** exhibited the largest divergence between sales and profits, with sales exceeding profits by **$9,994,346.21**.
3. **Profit Trends**:
   * Overall profits declined by **16.16%** between **2017 and 2019**.
   * The most significant decline occurred between **January 2017** and **October 2019**, where profits dropped by **21.15%**, from **$9,598,696.65** to **$7,568,565.85** during the steepest period.

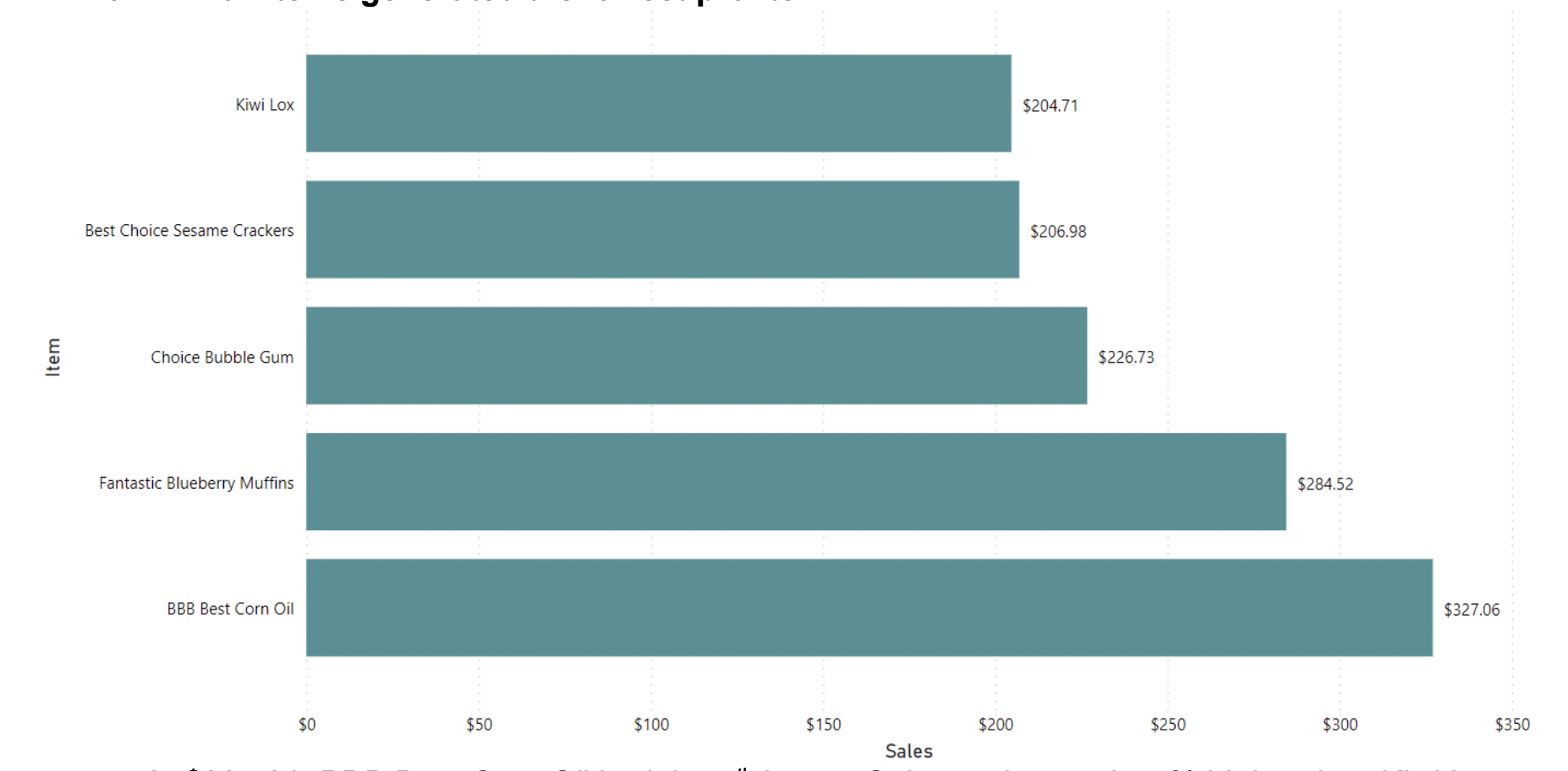
**5. How was the growth in Profit year over year?**



**Insights and Implications:**

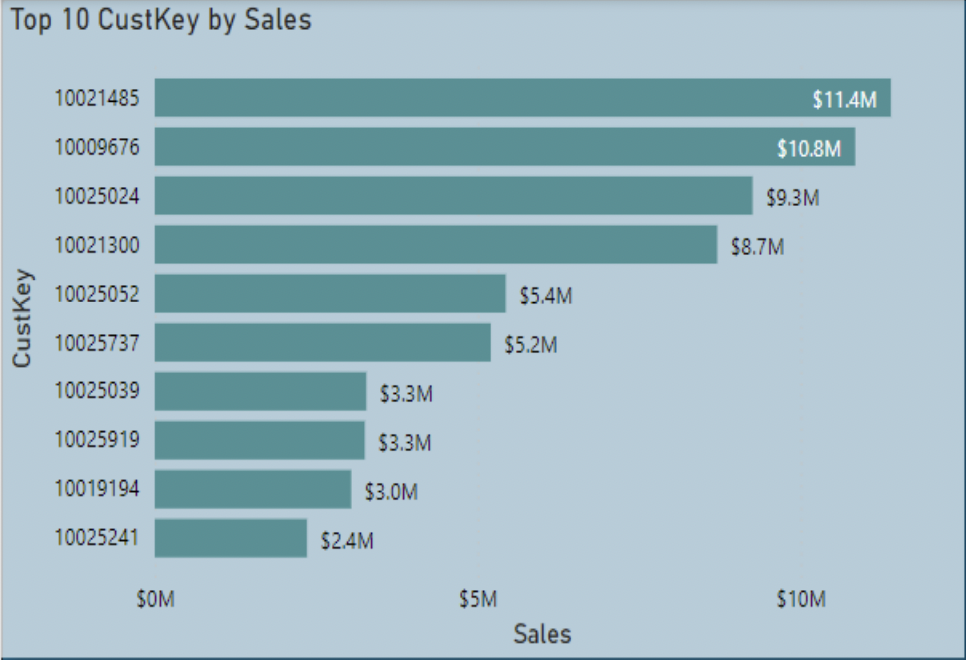
* **2018** represents a significant dip in profitability, requiring further investigation into operational inefficiencies, market dynamics, or external factors impacting sales.
* The recovery in **2019** underscores the effectiveness of corrective strategies or market recovery, but the gap from **2017** indicates room for further improvement.

**6.Which Items generated the lowest profits?**



**Low-Selling Items Analysis**

This bar chart showcases the **lowest-selling items** based on total sales:

1. **Top 5 Lowest-Selling Items**:
   * **Kiwi Lox**: Achieved the lowest sales at **$204.71**.
   * **Best Choice Sesame Crackers**: Recorded sales of **$206.98**, slightly higher than Kiwi Lox.
   * **Choice Bubble Gum**: Generated sales of **$226.73**.
   * **Fantastic Blueberry Muffins**: Slightly outperformed others with sales of **$284.52**.
   * **BBB Best Corn Oil**: Though still among the lowest, had the highest sales in this category at **$327.06**.
2. **Key Observations**:
   * These products are underperforming in terms of sales and may require strategic interventions such as targeted promotions, bundling, or inventory optimization.
   * The wide variation in sales among these items highlights potential gaps in demand or marketing focus.
3. Who were the Top 10 Customers?

**Key Observations:**

1. **Top Customers**:
   * **Customer 10021485**: The highest contributor to sales, with **$11.4M**.
   * **Customer 10009676**: A close second, contributing **$10.8M** in sales.
   * **Customer 10025024**: Ranked third, with **$9.3M** in sales.
2. **Mid-Tier Customers**:
   * Customers like **10021300** and **10025052** contributed **$8.7M** and **$5.4M**, respectively.
3. **Lower Tier**:
   * The bottom customers in the top 10, such as **10025241**, contributed **$2.4M**.

#### ****Link Between Quantitative Methods and Practical Insights****

* The applied **logistic regression analysis** provided actionable insights into the factors influencing sales likelihood:
  + Discounts should be used as a key lever to drive conversions, especially for price-sensitive customers.
  + Inventory management should prioritize high-demand products to maximize sales likelihood.
  + Pricing strategies should focus on combining attractive discounts with optimized list prices to enhance overall sales.
* These findings are directly actionable in marketing, pricing, and inventory decisions, enabling the development of targeted strategies to maximize sales and profitability.