



buncha



Customer & Operations Analysis

Insights, Visualizations, and Business Recommendations





Introduction

- **Overview:** Analyzing customer behavior, sales trends, and operational efficiency to derive actionable insights to enhance business performance, improve customer retention, and optimize inventory management.
- **Datasets used:** orders1.csv, customers.csv, inventory.csv, delivery_performance.csv, discount_campaign.csv, sales_data.csv
- **Tools & techniques applied:** SQL, MS Excel, Python, MySQL Workbench, Visual Studio Code



SQL Analysis – Customer Purchase & Delivery Trends

Key objectives: To understand customer retention, purchasing trends, and delivery performance



SQL Queries

- Customers who haven't placed an order in the last 60 days but had at least 2 orders before

```
SELECT DISTINCT customer_id FROM orders1
WHERE customer_id IN (
    SELECT customer_id FROM orders1
    GROUP BY customer_id
    HAVING COUNT(order_date) >= 2
)
GROUP BY customer_id
HAVING MAX(order_date) < DATE_SUB(CURDATE(), INTERVAL 60 DAY);
```



SQL Queries

- The average time between consecutive orders for repeat customers

```
FOR EACH CUSTOMER
WITH GAP_BETWEEN_ORDERS AS (
    SELECT customer_id,
           DATEDIFF(order_date, LAG(order_date) OVER (PARTITION BY customer_id ORDER BY order_date)) AS
           DAYS_BETWEEN_ORDERS
    FROM orders1
    WHERE customer_id IN (
        SELECT customer_id FROM orders1
        GROUP BY customer_id
        HAVING COUNT(order_date) > 1 -- repeat customers
    )
)
SELECT customer_id, AVG(DAYS_BETWEEN_ORDERS) AS avg_time_btwn_consecutive_orders_for_each_customer
FROM GAP_BETWEEN_ORDERS
WHERE DAYS_BETWEEN_ORDERS IS NOT NULL
GROUP BY customer_id;
```



SQL Queries

- Top 10% of customers by total spend and their average order value

```
WITH customer_spending AS (  
    SELECT customer_id, SUM(total_amount) AS total_spend,  
    COUNT(ORDER_DATE) AS NO_OF_ORDERS  
    FROM orders1  
    GROUP BY customer_id),  
ranked_customers AS (  
    SELECT customer_id, total_spend, TOTAL_SPEND/NO_OF_ORDERS AS AOV,  
    PERCENT_RANK() OVER (ORDER BY total_spend DESC) AS percentile_rank  
    FROM customer_spending)  
SELECT customer_id, total_spend, AOV  
FROM ranked_customers  
WHERE percentile_rank <= 0.1 -- only top 10%  
ORDER BY total_spend DESC;
```



SQL Queries

- Delivery time efficiency by calculating the percentage of on-time deliveries per region

```
WITH delivery_times AS (  
    SELECT city, COUNT(CASE WHEN delivery_status = 'On Time' THEN 1 END) AS on_time_deliveries,  
           COUNT(*) as total_deliveries  
    FROM ORDERS1 ORD INNER JOIN DELIVERY_PERFORMANCE DP  
    ON ORD.ORDER_ID = DP.ORDER_ID  
    GROUP BY city)  
SELECT city, on_time_deliveries*100/total_deliveries AS delivery_time_efficiency  
FROM delivery_times  
GROUP BY city  
ORDER BY delivery_time_efficiency DESC;
```

Key Findings

- Customer IDs of customers who haven't placed an order in the last 60 days but had at least 2 orders before are – 1102, 1045, 1301, 1252, 1443, and 959 others.

	customer_id
▶	1102
	1045
	1301
	1252
	1443
	1335
	1262
	1802
	1730
	1194
	1683
	1241
	1914
	1707
	1614
	1988
	1310
	1885
	1261
	1323
	1494
	1725
	1995

Key Findings

- The average time between consecutive orders for repeat customers are different for different customers, but the average time for all the customers is approximately 67 days.

	customer_id	avg_time_btwn_consecutive_orders_for_each_customer
▶	1001	45.6667
	1002	225.0000
	1003	141.5000
	1005	24.5000
	1006	95.6667
	1007	133.5000
	1008	109.0000
	1009	30.3000
	1010	28.2857
	1011	49.6000
	1012	35.6667
	1013	78.3333
	1014	46.6667
	1015	54.0000
	1016	50.0000
	1017	32.5714
	1018	60.0000
	1019	61.4000
	1020	48.5000
	1021	88.3333
	1022	33.0000
	1023	44.4286
	1024	91.5000

Key Findings

- The top 10% of customers by total spend and their average order value are shown in the table.
- Only 100 customers constitute the top 10% of customers by total spend and their AOV.

customer_id	total_spend	AOV
1161	3814.541692994909	293.4262840765315
1770	3528.258650383805	320.75078639852774
1578	3473.9336559645362	315.81215054223054
1200	3331.4805743765755	302.8618703978705
1916	3161.905978828656	351.32288653651733
1118	3059.8541512909915	305.9854151290991
1661	3055.111883200501	339.4568759111668
1678	3031.608960707822	303.1608960707822
1304	3027.0228225517835	336.3358691724204
1806	3016.592373085105	301.6592373085105
1685	2991.1652940324684	299.11652940324683
1017	2986.732973099488	373.341621637436
1672	2956.4824401708765	328.49804890787516
1359	2915.627352300833	323.9585947000926
1473	2900.802092447907	322.311343605323
1238	2880.9706167629124	360.12132709536405
1280	2856.124781629137	317.347197958793
1521	2854.5134140364253	285.4513414036425
1798	2854.3426939553065	317.14918821725627
1063	2809.0397254954723	280.90397254954723
1164	2785.5197464086805	348.18996830108506
1243	2782.600366839609	347.8250458549511
1788	2758.2946956800633	250.75406324364212



Key Findings

- The percentage of on-time deliveries per region is shown in the table.
- It is observed that Houston has the most efficient delivery service as 84.9% of the deliveries were on time.

city	delivery_time_efficiency
Houston	84.9076
New York	84.5766
San Francisco	84.5420
Los Angeles	84.3040
Chicago	83.6588



Excel Analysis – Sales & Inventory Dashboard

Key objectives: To analyze sales, inventory levels, and out-of-stock patterns

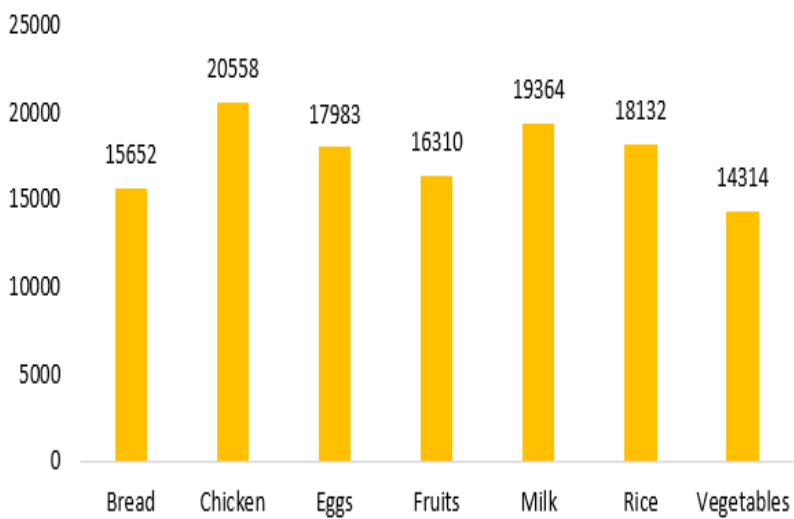


Sales and Inventory Dashboard

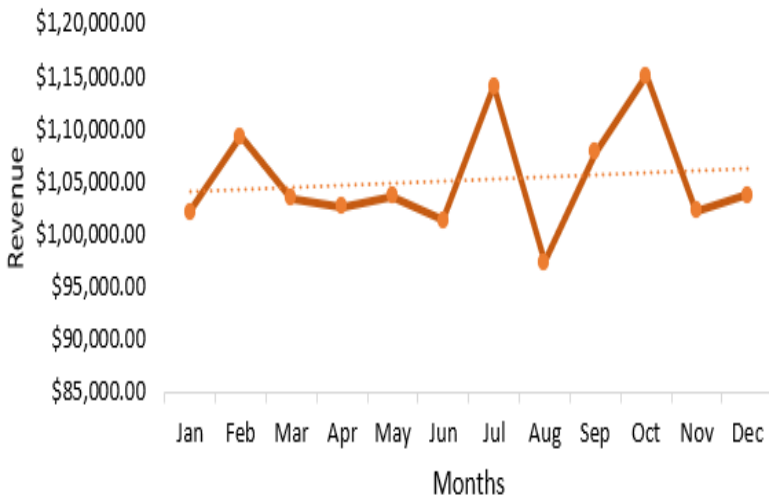
CRR: 36.81%

AOV: \$ 252.57

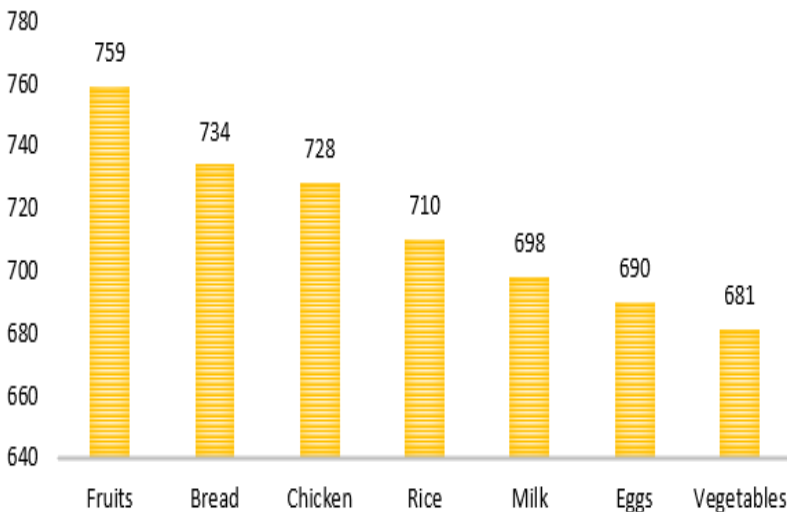
Stock Levels



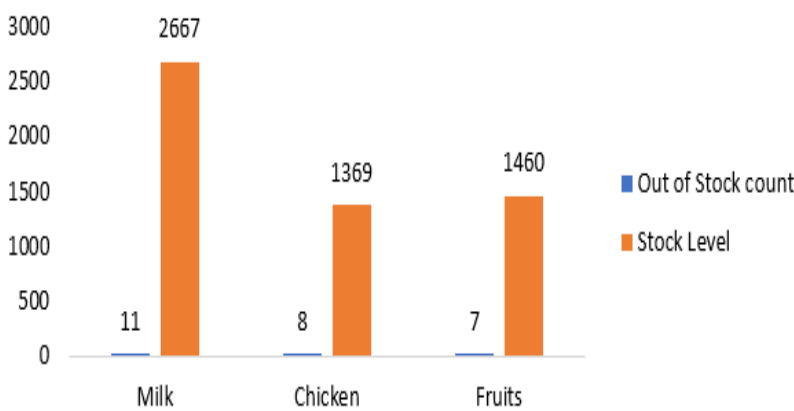
Monthly Revenue Growth



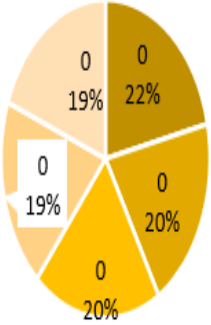
PRODUCT DEMAND



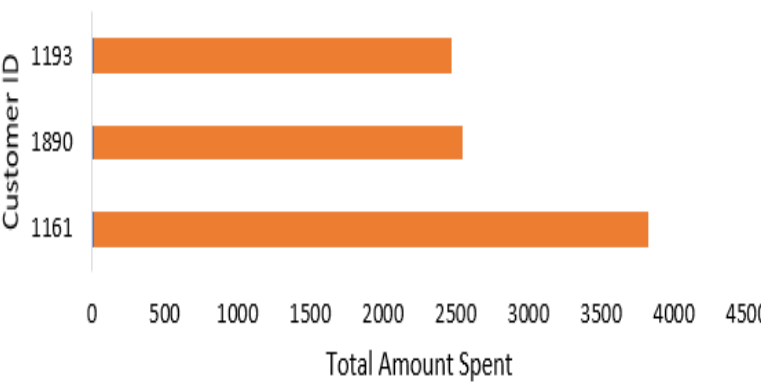
Top 3 products frequently getting Out of Stock



City-wise Revenue Contribution



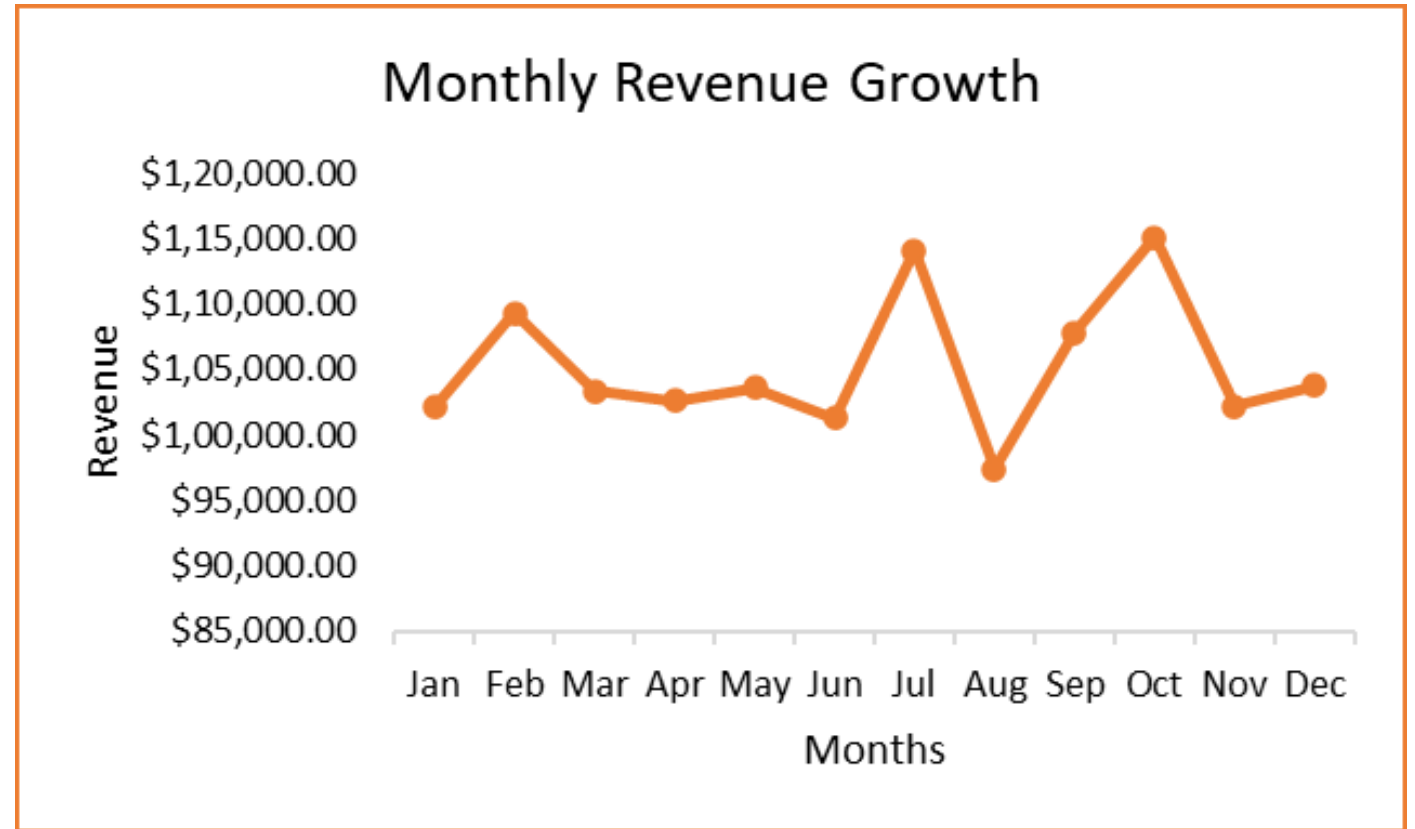
Top 3 Repeating Customers vs Revenue



Monthly Revenue Growth

Key Findings:

- The revenue shows significant ups and downs, indicating seasonal and campaign-driven variations.
- February, July, September, and October see noticeable peaks.
- There is a sharp decline in August and November, highlighting potential issues like low demand, or stock shortages.
- March to June shows relatively stable revenue, suggesting a consistent sales phase without major disruptions.





Average Order Value

Key Findings:

- The Average Order Value (AOV) is calculated using the formula:
$$AOV = \text{Total Revenue} / \text{Number of Orders}$$
- Total Revenue = \$1,262,860.39
- Number of Orders = 5000
$$AOV = 1,262,860.39 / 5000$$
$$= 252.57$$
- **Average Order Value (AOV):** \$ 252.57 per order, suggesting high-value purchases per transaction.
- Increasing the number of orders while maintaining AOV can drive higher revenue.



Customer Retention Rate

Key Findings:

- The Average Order Value (AOV) is calculated using the formula:

$$CRR = \left(\frac{\text{Customers at End of Year} - \text{New Customers Acquired}}{\text{Customers at Start of Year}} \right) \times 100$$

- No. of unique customer IDs who made a purchase in January = 345
- No. of unique customer IDs who made a purchase in December = 355
- No. of customers who made their first purchase within the year (but were not in the start-of-year list) = 228
- Therefore,

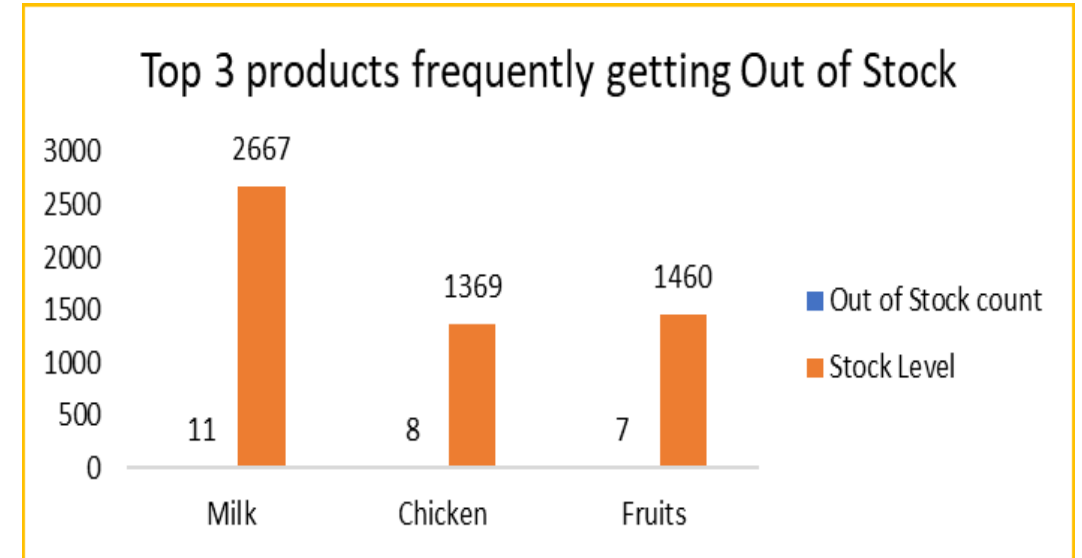
$$\begin{aligned} CRR &= \{(355 - 228) / 5000\} * 100 \\ &= 36.81\% \end{aligned}$$

- Customer Retention Rate(CRR):** CRR 36.81% indicates that more than 60% of customers did not return after their initial purchase.
- It suggests for a need for better retention strategies.

Products frequently out of stock

Key Findings:

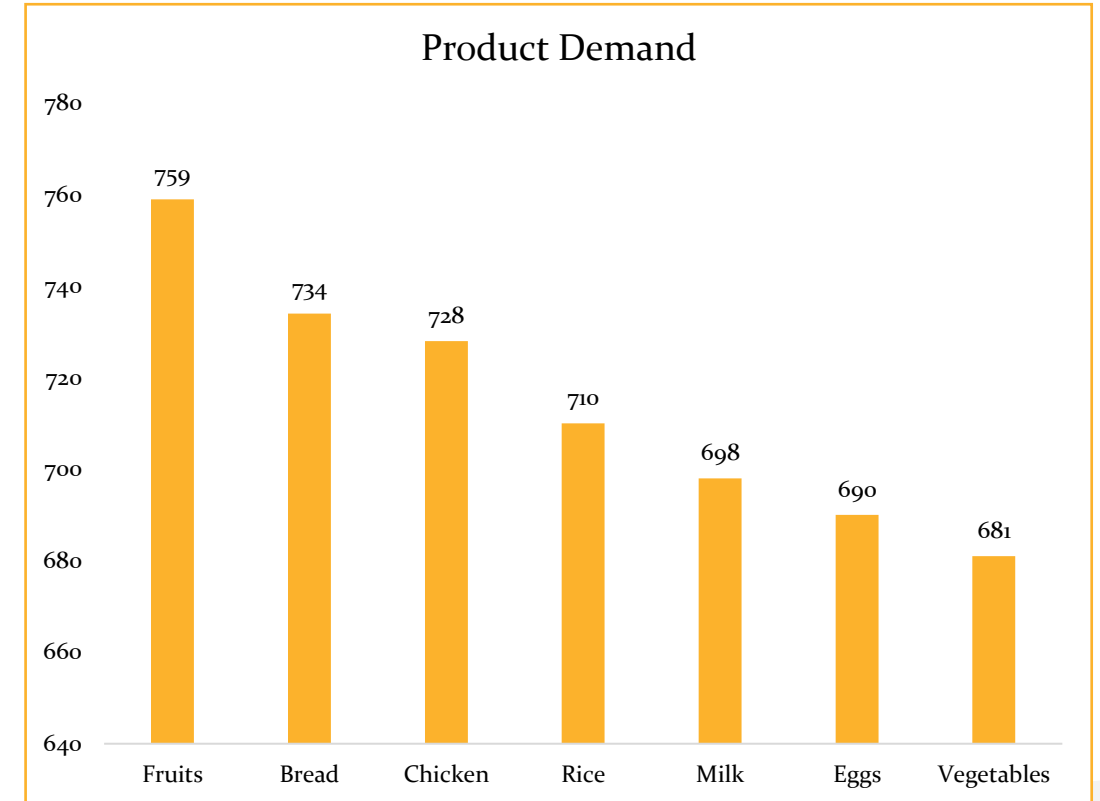
- Top 3 products that frequent get out of stock are – milk, chicken and fruits.
- Milk has the highest out-of-stock issue – It was out of stock 11 times, with a stock level of 2667, indicating high demand and possible supply chain inefficiencies.
- Chicken follows with 8 stock-outs – Despite having a stock level of 1369, it faced frequent stock-outs, suggesting inventory mismanagement.
- Fruits were out of stock 7 times – With a stock level of 1460, fruits also faced availability issues, which might impact sales and customer satisfaction.



Product Demand

Key Findings:

- Fruits have the highest demand (759 units), suggesting they are a key product category, requiring efficient stock management to avoid shortages.
- Bread and Chicken follow closely with 734 and 728 units sold respectively, contributing significantly to overall sales.
- Rice, Milk, and Eggs show moderate demand with their sales range between 690-710 units, indicating they are steady-selling products.
- Vegetables have the lowest demand (681 units) indicating lower customer preference or seasonal variations in consumption.
- Business Impact – High-demand products like Fruits, Bread, and Chicken should be prioritized restocking, while lower-demand items might require targeted promotions to boost sales.





Python Analysis – Customer Segmentation & Demand Patterns

Key objectives: To segment customers based on purchasing behavior and detect demand trends



Methodology: K-Means Clustering for Customer Segmentation

Step 1: Data Preprocessing

- Loaded datasets and converted order_date to datetime format.
- Extracted month names for trend analysis.

Step 2: Customer Segmentation (K-Means)

- Scaled total_spend and num_orders.
- Used Elbow Method to determine K=4.
- Applied K-Means clustering and visualized segments.

Step 3: Sales Trend Analysis

- Aggregated monthly sales and sorted by calendar order.
- Plotted revenue trends with data labels.

Step 4: Customer Spending Behavior

- Grouped total spending per customer and merged with segments.
- Visualized spending patterns using a boxplot.

Step 5: Order Distribution by City

- Counted total orders per city.
- Visualized order distribution with a bar chart and data labels.

Customer Segmentation Based on Spending and Order Frequency

Insights:

Four Distinct Customer Segments:

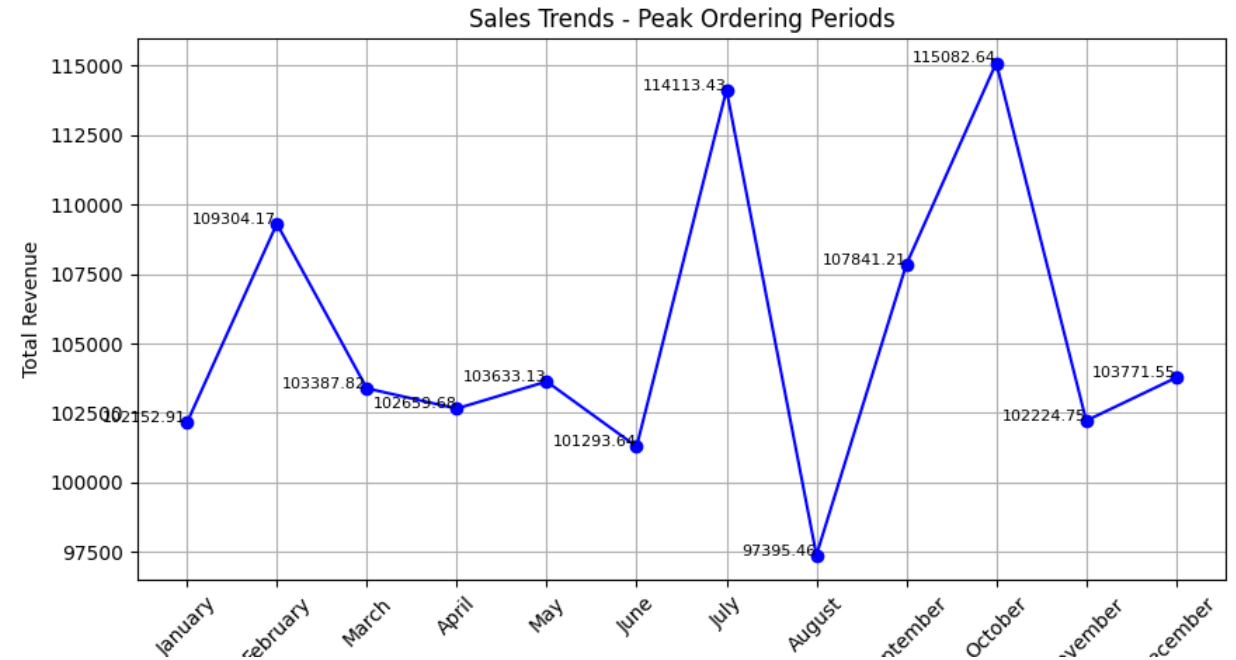
- **Red Cluster (High Spend, High Orders, High Value Buyers):** Most valuable customers, high engagement.
- **Green Cluster (Low Spend, High Orders, Frequent Buyers):** Frequent shoppers but lower spending per order.
- **Blue Cluster (Low Spend, Low Orders, Occasional Buyers):** Low-range customers with lower spending.
- **Purple Cluster (High Spend, Low Orders):** Occasional big spenders, potential for loyalty programs.



Sales Trends – Peak Ordering Periods

Insights:

- **Peak Revenue Months:** July (USD 114,113.43) and October (USD 115,082.64) saw the highest sales.
- Revenue varies month-to-month, suggesting external factors like discounts, holidays, or economic conditions influence sales.
- Boost marketing in July and October to maximize sales.



Customer Spending Patterns by Segment

Insights:

- All customer segments have comparable median spending, indicating no significant difference in typical spending behavior.
- Each segment has a broad range of spending, with minimums close to zero and maximums exceeding 3,000, suggesting a mix of low and high-value customers.
- Some customers in each segment spend significantly more than others, which could indicate premium buyers or bulk purchasers.



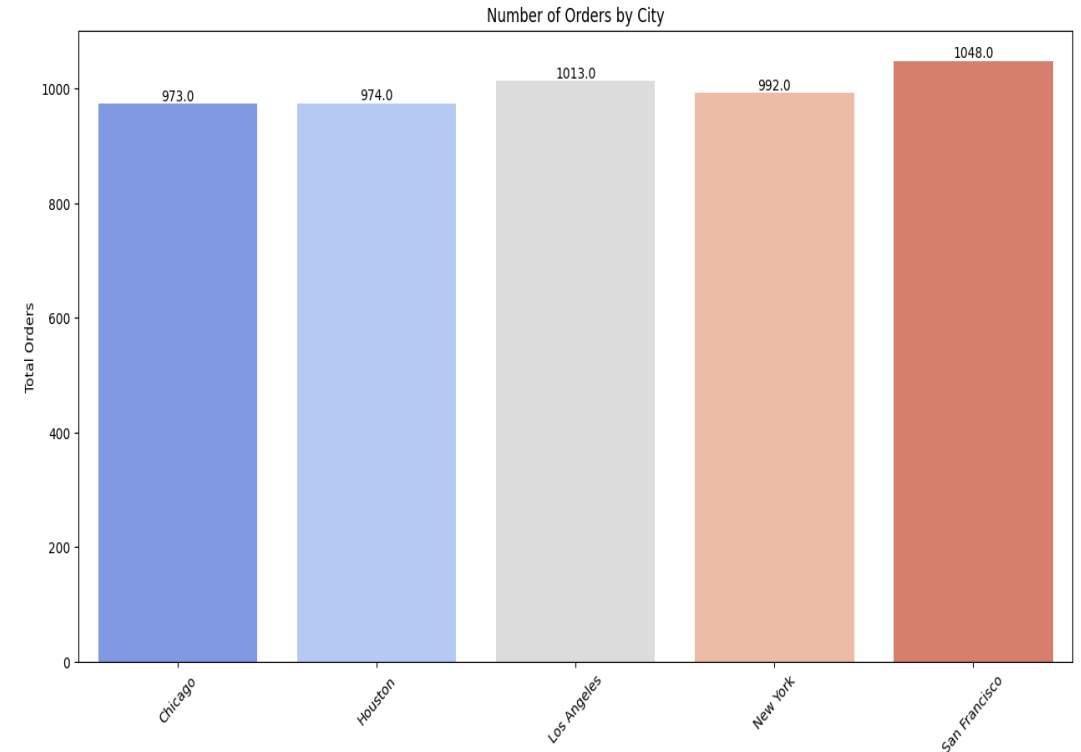
City-wise Order Count

Insights:

- San Francisco has the highest number of orders (1,048), suggesting strong customer engagement in this city.
- Los Angeles follows closely with 1,013 orders, indicating high demand.
- Chicago and Houston have the lowest number of orders (973 and 974), but the difference is not significant.
- New York is in the mid-range with 992 orders, showing steady performance.

Recommendations:

- Focus marketing efforts in San Francisco and Los Angeles to capitalize on existing demand.
- Explore reasons for slightly lower orders in Chicago and Houston, such as competition or customer preferences.
- Maintain consistent strategies in New York to sustain order volume.





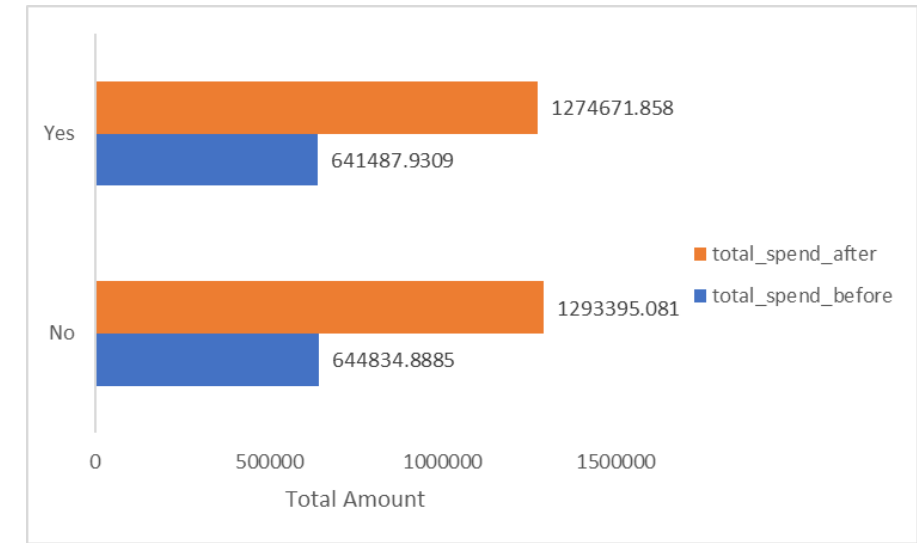
Business Case Study– Discount Impact Analysis

Key objectives: To analyze the impact of discount strategies on profitability and retention

Pre vs post-discount customer spending behavior

Key Findings:

- Total spend before discounts: 1,286,322.82
- Total spend after discounts: 2,568,066.94
- Customers who did not apply discounts spent 644,834.89 before and 1,293,395.08 after.
- Customers who applied discounts spent 641,487.93 before and 1,274,671.86 after.
- Overall, spending increased significantly after the discount campaign, but both groups showed similar trends.
- This suggests that the spending boost was not entirely driven by the discount campaign—other factors might have influenced it, such as seasonal demand, marketing efforts, or natural customer behavior.



Did the Discount Campaign Work?

- If the goal was to **increase overall spending**, then yes, spending did go up.
- However, if the goal was to **specifically drive more spending among those who applied discounts**, the impact is unclear because even non-discounted customers increased their spending.



Insights on Customer Responsiveness to Discounts

Key Findings

- Highly Responsive (263 customers) → These customers significantly increased their spending after receiving discounts, making them the most valuable group for future promotions.
- Moderately Responsive (82 customers) → Showed some increase in spending, but not as dramatically. They might need better-targeted offers.
- Unresponsive (25 customers) → Their spending behavior remained unchanged despite the discount. Discounts may not be the right incentive for them.
- Negative Response (130 customers) → These customers reduced their spending even after receiving discounts. This could indicate they are price-sensitive and may have stocked up earlier or shifted to a competitor.

Key Takeaways


- Discounts worked well for a large portion ($263 + 82 = 345$ customers).
- A significant 130 customers reduced spending, which needs further investigation.



Strategy to Maximize Revenue While Maintaining Profitability

Key Findings

- Not all customers respond positively to discounts – some increased their spending, while others reduced orders.
- Highly responsive customers increased both order count and spending significantly.
- Some customers spent less despite receiving discounts, which indicates ineffective discount allocation.
- Customers who did not receive discounts but still increased spending may be naturally high-value customers.



Strategy to Maximize Revenue While Maintaining Profitability

Recommendations

1. Targeted Discounts

- Offer tiered discounts (e.g., 10% for moderate, 15% for high-value buyers). Use cashback or free shipping for unresponsive customers instead of price cuts.
- Avoid discounts for negative response customers; focus on retention (bundles, subscriptions).

2. Strategic Timing

- Use limited-time, seasonal, and event-based discounts instead of constant offers.

3. Maximize Order Value

- Set minimum spend thresholds (e.g., "10% off on orders above \$500").
- Upsell & cross-sell (e.g., "Buy X, get Y at 20% off").



Potential Future Improvements

- Implement predictive analytics for demand forecasting to reduce stock-outs.
- Enhance customer retention strategies through targeted promotions and loyalty programs.
- Optimize pricing and discount strategies based on segment behavior.
- Automate real-time dashboards for continuous performance monitoring.



Thank You!