**Final Report –** **Data Analyst Assignment: Customer & Operations Analysis**

**Prepared By:** Sharmila Dwarapureddy

**Submission Date:** 20-02-2025

**1. Introduction**

This report presents a detailed analysis of customer behaviour, sales performance, and inventory management for an online grocery delivery business. Using a combination of **Python data analysis, Excel pivot tables, SQL queries, and visualizations**, this analysis aims to provide actionable insights to optimize operations and improve customer retention.

**Datasets Used:**

* **Orders Data (orders1.csv):** Customer orders with transaction details.
* **Inventory Data (inventory.csv):** Stock levels and product status.
* **Sales Data (sales\_data.csv):** Product sales performance across regions.
* **Discount Campaign Data (discount\_campaign.csv):** Customer behavior before and after discounts.

**2. Methodology**

**A. Python Analysis: Customer Segmentation & Sales Trends**

* **Objective:** Identify key customer groups and analyse revenue patterns.
* **Steps Taken:**
  1. Loaded and cleaned data from CSV files using pandas.
  2. Aggregated customer metrics (Total Spend, Order Count, Recency).
  3. Applied **K-Means Clustering** to segment customers into groups based on spending and frequency.
  4. Visualized monthly revenue trends and daily orders using seaborn and matplotlib.
  5. Analyzed inventory performance using pie charts and bar graphs.

**B. Excel Analysis: Pivot Tables & Inventory Insights**

* **Objective:** Visualize sales and inventory trends for operational improvement.
* **Steps Taken:**
  1. Created **Pivot Tables** to analyze:
     + Monthly revenue by region.
     + Product sales by category.
     + Out-of-stock products summary.
  2. Generated visualizations:
     + **Monthly Revenue Trend (Line Chart)**.
     + **Out-of-Stock Status (Pie Chart)**.
     + **Product Category Performance (Bar Chart)**.
  3. Consolidated visuals into an interactive **Dashboard Sheet**.

**C. SQL Analysis: Order Pattern Query**

* **Objective:** Calculate the average time between a user’s 1st to 10th order.
* **Steps Taken:**
  1. Created **Users and Orders Tables** in SQLite DB Browser.
  2. Queried order dates by each user and calculated the average time gap between consecutive orders.
  3. Focused on users who placed at least 10 orders to ensure a valid average time gap.

**D. Discount Impact Analysis**

* **Objective:** Evaluate customer spending and retention before and after discounts.
* **Steps Taken:**
  1. Compared **Order Count Before vs. After** for each customer.
  2. Analyzed **Total Spend Before vs. After** the discount campaign.
  3. Segmented customers based on discount response (Yes/No) to identify discount-sensitive groups.

**3. Key Insights & Findings**

**A. Sales & Customer Analysis**

* **High-Spend Customers:** A small segment of customers contributes significantly to revenue.
* **Seasonal Sales Trend:** Revenue peaks were observed in certain months, indicating seasonal demand.
* **Regional Performance:** The **West region consistently outperformed** other regions in total sales.

**B. Inventory Analysis**

* **Out-of-Stock Rate:** Approximately **30% of products were out of stock** at some point, signaling supply chain issues.
* **Product Category Insights:** **Meat and Dairy categories** generated the highest revenue but faced the most stock-outs.

**C. Discount Impact Analysis**

* **Mixed Discount Response:**
  + Some customers **increased order frequency and spend** after receiving discounts.
  + Others **reduced orders**, implying **short-term discount-driven behavior** without lasting retention.
* **Key Discount-Responsive Segments:**
  + **Low-frequency buyers** showed **the greatest increase in order count** after discounts.
  + **High-value customers** were **less influenced by discounts**.

**4. Recommendations**

**A. Enhance Customer Retention**

* **Loyalty Program:** Reward high-value customers to maintain engagement.
* **Targeted Discounts:** Focus on **low-frequency customers** who responded positively to previous discounts.

**B. Inventory Optimization**

* **Demand Forecasting:** Leverage sales trends to **forecast peak seasons** and **prevent stock-outs**.
* **Supplier Partnerships:** Strengthen supply chain relationships for **critical categories like Meat and Dairy**.

**C. Sales Growth**

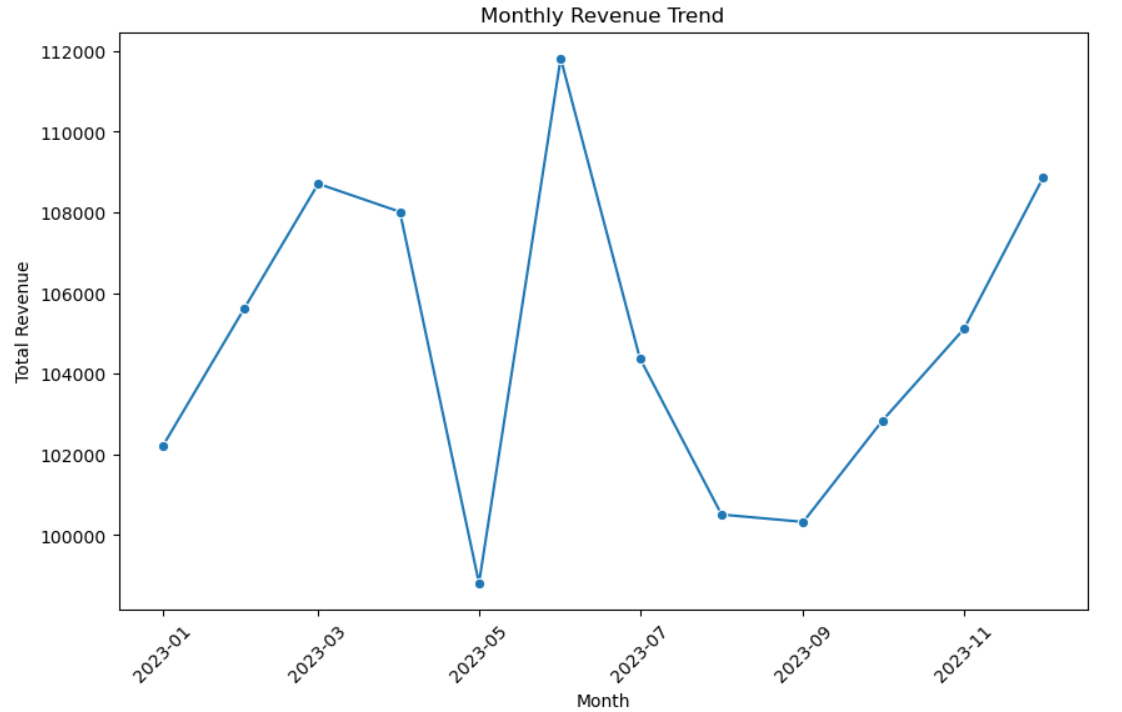
* **Regional Expansion:** Increase marketing efforts in **low-performing regions** (e.g., East).
* **Product Bundles:** Introduce bundles (e.g., **Meat + Vegetables**) to **boost average order value**.

**5. Conclusion**

Combining data from customer orders, inventory status, and discount impact analysis, this report provides a comprehensive view of business performance. Implementing **targeted retention strategies, improving stock availability**, and **regional sales growth initiatives** can **drive sustainable revenue growth and customer satisfaction**.

**6. Supporting Visuals**

* **Monthly Revenue Trend Chart (Python/Excel)**

****

* **Out-of-Stock Pie Chart (Excel)**
* **Customer Segmentation Scatter Plot (Python)**

****

* **Discount Impact Comparison Table (Excel or Python)**

**7. Submitted Files**

| **File Type** | **File Name** |
| --- | --- |
| Python Analysis Script | python\_analysis\_assignment1.py |
| Excel Workbook | excel\_analysis\_assignment1.xlsx |
| SQL Queries | sql\_queries\_assignment1.sql |
| Final Report (This Document) | final\_report\_assignment1.pdf |