AMAZON SALES

Prepared by: Geoffrey Otobor Onajefe

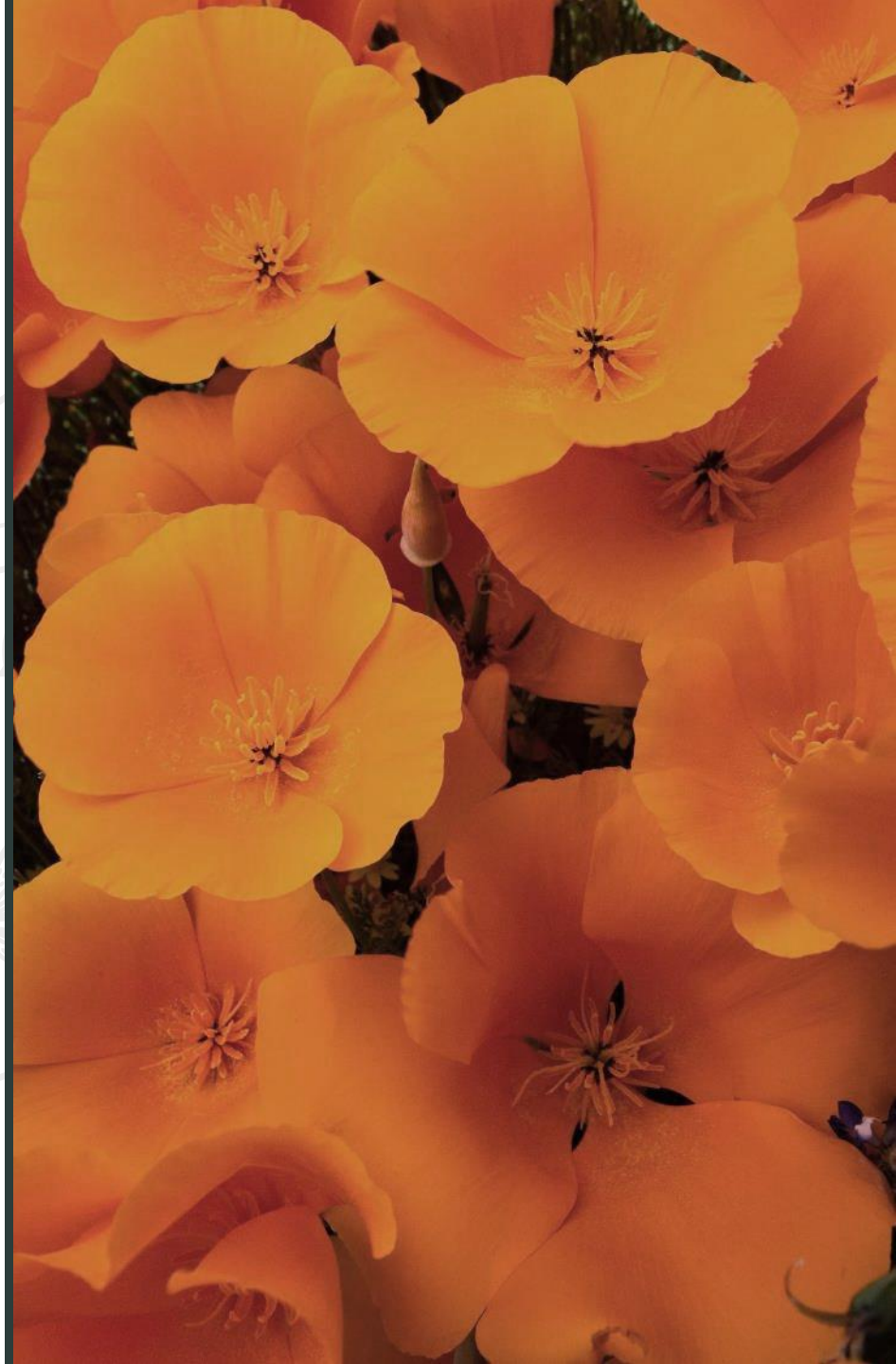
Date: 14th June 2025

Introduction

The Amazon sales data Analysis project aims to explore and analyze sales data from Amazon to uncover key trends, patterns, and insights that can drive business decisions. The dataset includes information such as order details, product categories, customer demographics, payment methods, and sales performance. By leveraging Python and data visualization tools like Pandas and Matplotlib, this project seeks to transform raw sales data into actionable intelligence.

OBJECTIVES

- **1. Data Exploration & Cleaning**
- **2. Sales Performance Analysis**
- **3.Customer & Payment Insights.**
- **4. Visualization & Reporting**



DATA OVERVIEW

- **Source**: Simulated Amazon Sales Data (Kaggle 2025)
- **Total Columns**: 11
- **Sample Size**: **Size**: 250 rows (representing 250 individual orders)

Column Name	Description
Order ID	Unique identifier for each order
Date	Date when the order was placed
Product	Name of the item purchased
Category	Product category (e.g., Electronics, Clothing)
Price	Price per unit
Quantity	Quantity purchased
Total Sales	Revenue per order (Price × Quantity)
Customer Name	Name of the buyer
Customer Location	Customer's city
Payment Method	Payment method used (e.g., PayPal, Credit Card)
Status	Status of the order (Completed, Cancelled, etc.)

This dataset provides a compact but rich view of sales transactions, allowing detailed analysis of trends, customer behavior, and business performance

METHODOLOGY

1. Data Collection & Preparation

Source**:** Simulated Amazon sales dataset (2025) containing 250 orders across 11 columns (e.g., product categories, payment methods, order status).

2. Data Cleaning

Handling Missing Values**:** Identified and addressed null entries (e.g., empty customer demographics) by either filling with defaults (e.g., "Unknown") or removing irrelevant rows.

Outlier Treatment: Used descriptive statistics (e.g., IQR for numerical fields like sales revenue) to detect and cap extreme values that could skew analysis.

Standardization**:** Normalized inconsistent formats (e.g., payment method spellings, date formats) to ensure uniformity.

METHODOLOGY (Part 2)

3. Exploratory Data Analysis (EDA)

Descriptive Statistics**:** Analyzed central tendencies (mean, median) and distributions of key variables (e.g., sales by category).

Trend Analysis: Used time-series decomposition (via line charts) to identify seasonal spikes (e.g., the Feb 16 sales peak).

4. Visualization & Reporting

Generated interactive charts (Matplotlib/Seaborn) to highlight trends, such as:

- Revenue concentration in electronics (bar charts).
- Payment method preferences (pie charts).
- Order fulfillment issues (heatmaps for cancellations).

Insights were synthesized into strategic recommendations (e.g., prioritizing high-margin products, improving logistics for cancellations).

METHODOLOGY (PART 3)

5. Tools used

- Python (Pandas, Matplotlib, Seaborn)
- Jupyter Notebook
- Excel

VISUALIZATION

•**Total sales by product (Bar Chart)**
Shows the total recorded by each product 2025

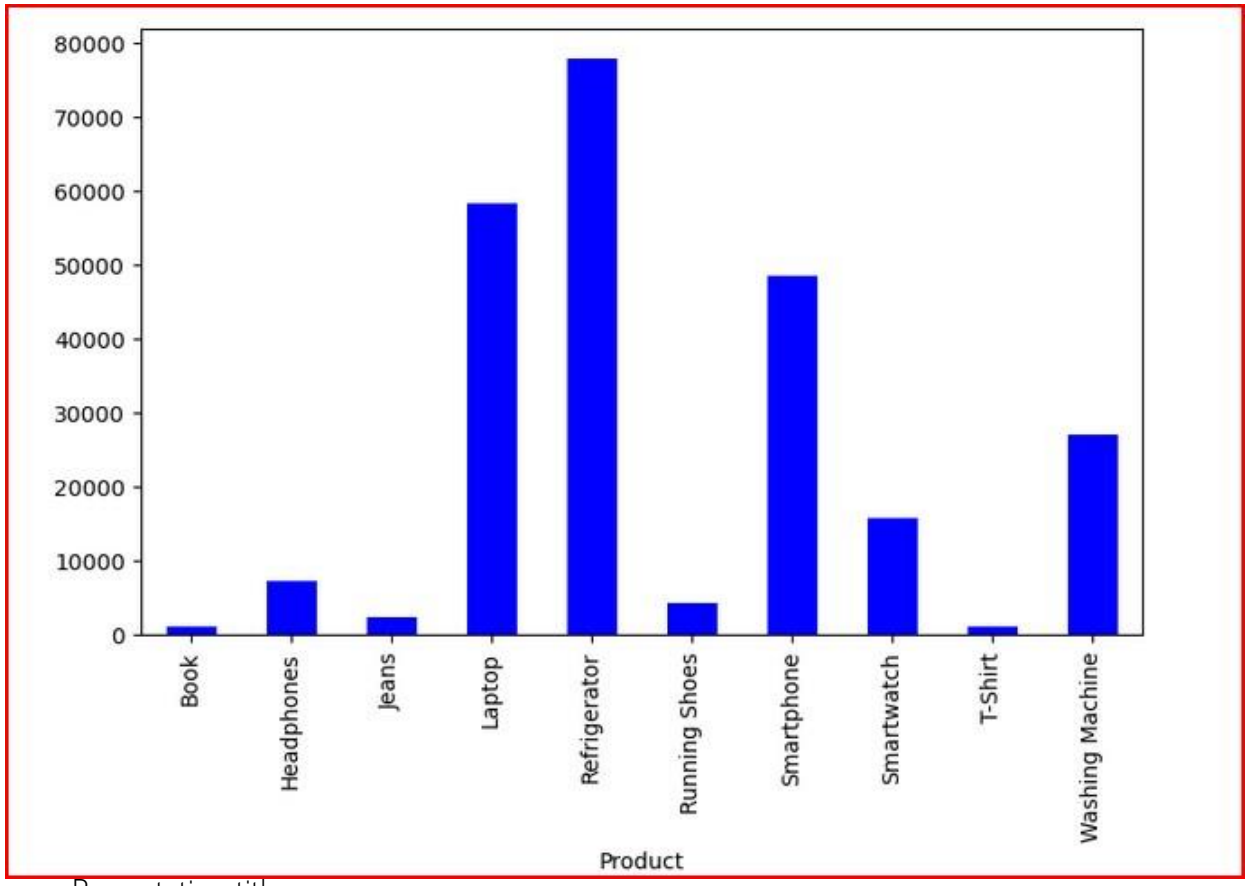
• **Order distribution by payment method (Pie Chart)**
Identified the percentage of each payment method

• **Total sales by date (Line Chart)**
Revealed the top ten dates with the highest sales

•**Compare average sales per product (horizontal bar chart)**
Highlighted the average sales per product

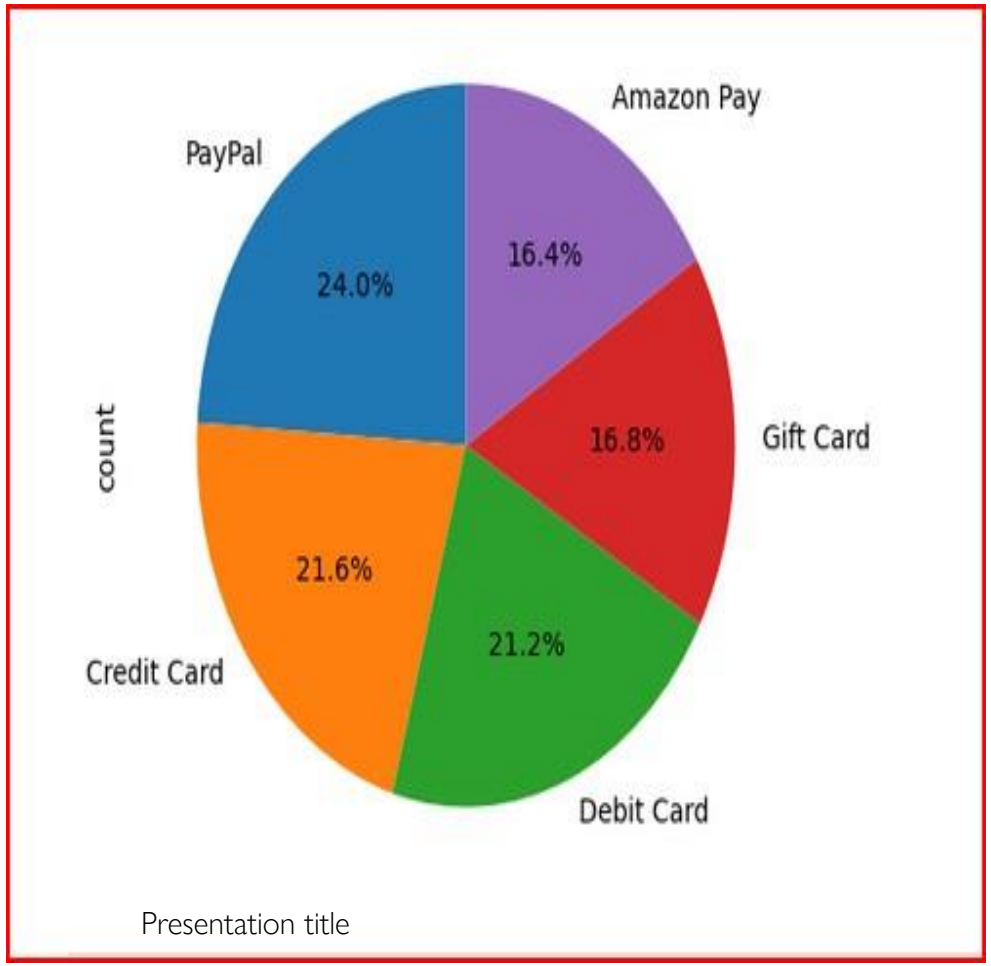
•**Order Status by category (Heat map)**
Identifies which category has the highest and lowest values

TOTAL SALES BY PRODUCT



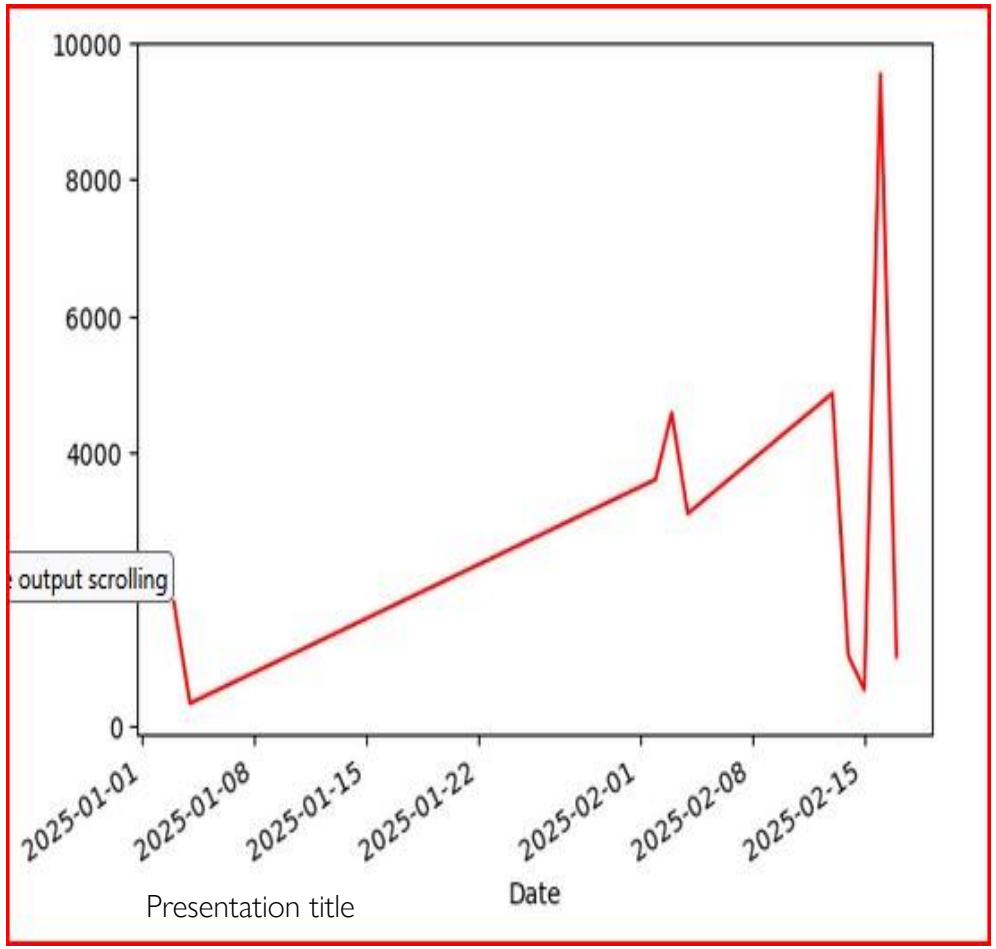
Observation: Refrigerators led sales with $78,000 in revenue, significantly outperforming laptops at $58,400 and smartphones at $48,500. In contrast, books, T-shirts, and jeans contributed modestly, ranging from $1,035 to $2,480, reflecting a clear consumer preference for high-value electronic products.

TOTAL ORDER BY PAYMENT METHOD



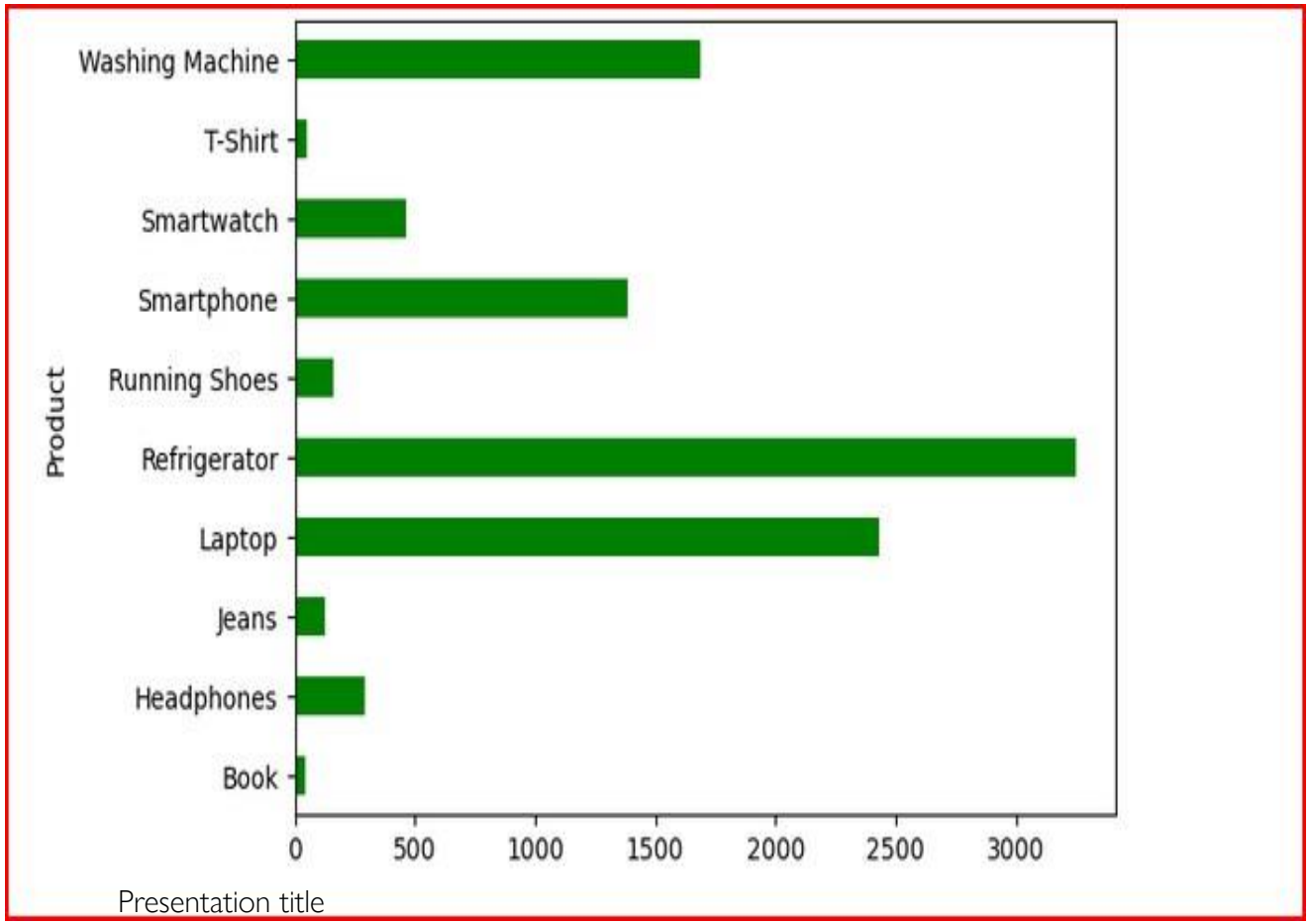
Observation: PayPal emerged as the most preferred payment method, accounting for 24% of orders, followed by Credit Card at 21.6% . In contrast, Amazon Pay and Gift Cards, with 16.4% and 16.8% respectively, were the least utilized, indicating lower customer preference for these options compared to digital wallets

TOP TEN SALES BY DATE



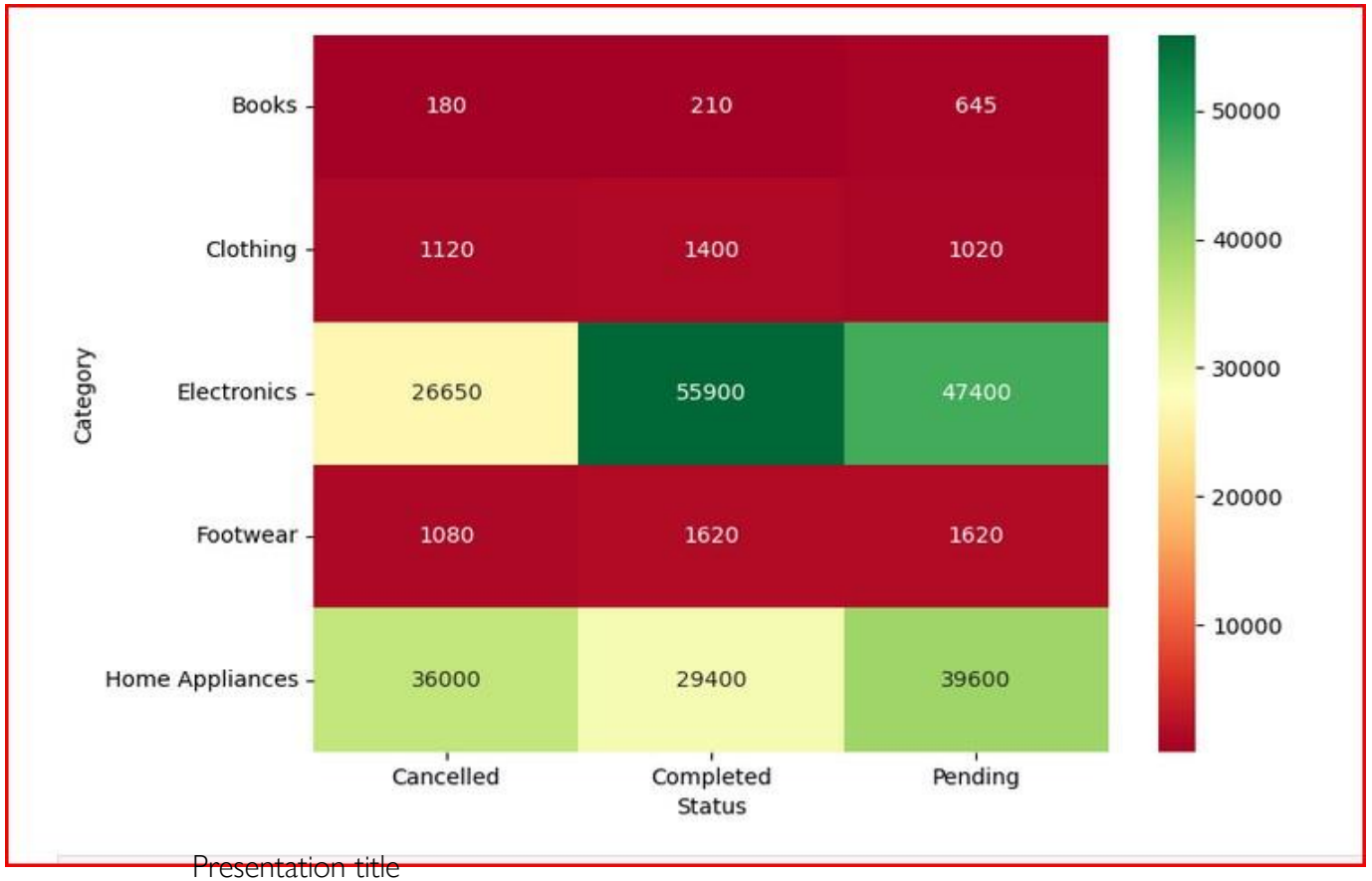
Observation: Sales reached their highest point on February 16, 2025, with a remarkable $9,540 in revenue, followed by a secondary peak on February 13 at $4,860. In contrast, the lowest revenue figures were observed on January 4 and February 15, amounting to $320 and $520 respectively. The dramatic surge in sales immediately after February 15 suggests the influence of a targeted event or promotional campaign that significantly boosted consumer activity on the 16th.

AVERAGE SALES BY PRODUCT



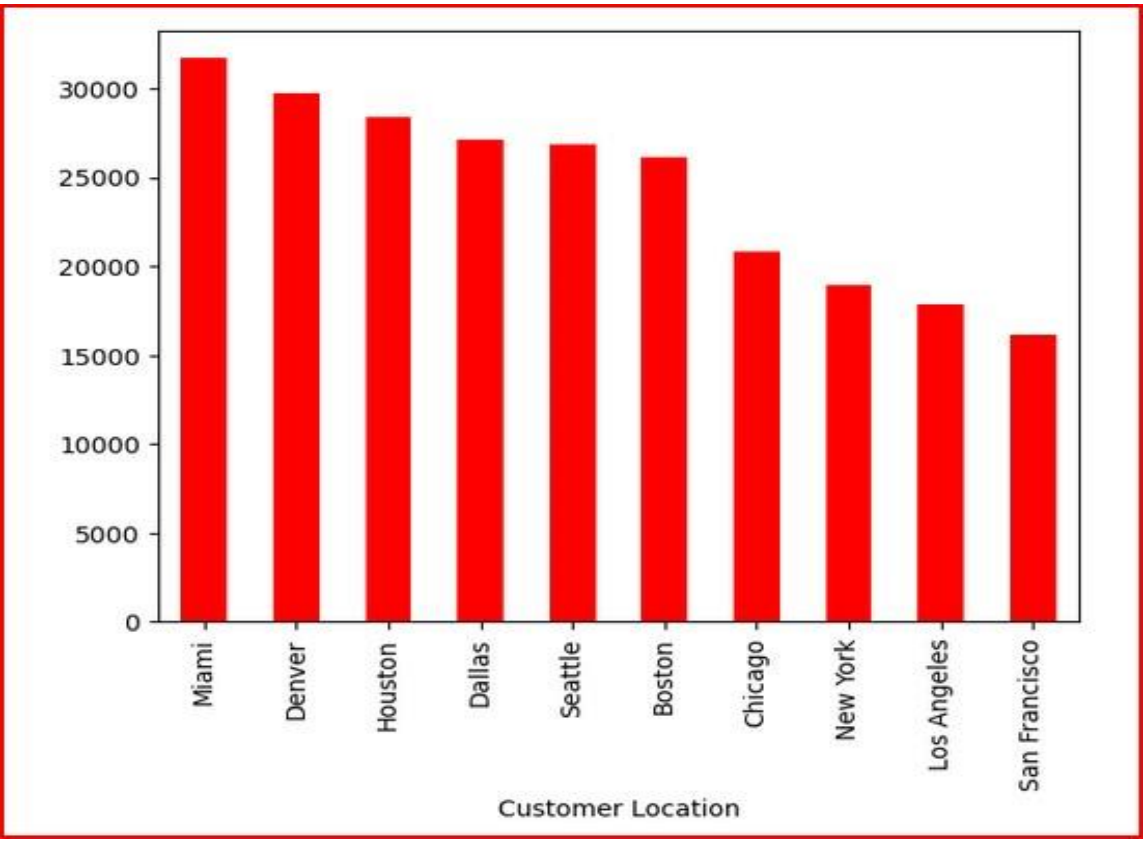
Observation. Refrigerators, Laptops, and Washing Machines generated the highest average revenues with $3,250, $2,433, and $1,687 respectively, highlighting strong profitability associated with high-value, durable goods. Conversely, Books, T-shirts, and Jeans recorded the lowest average revenues at $41.40, $53, and $124, indicating more modest consumer spending in these categories. This contrast underscores a clear consumer preference for premium, long-lasting products, offering strategic direction for businesses to prioritize marketing efforts and inventory planning toward these high-margin segments..

ORDER STATUS BY CATEGORY



Observation: Electronics recorded the highest figures in both completed and pending orders, with revenues of $55,900 and $47,400 respectively, followed by Home Appliances with $29,400 in completed orders. However, Home Appliances also experienced the highest order cancellations, totaling $36,000, with Electronics close behind at $26,650. While these categories demonstrate strong sales performance, the elevated cancellation rates indicate underlying issues in fulfillment or customer experience that warrant strategic attention to enhance reliability and satisfaction.

TOP TEN LOCATIONS BY SALES



Observation: Miami recorded the highest sales at $31,700, followed closely by Denver and Houston, indicating strong performance in southern and central regions. Coastal cities like New York, Los Angeles, and San Francisco trailed behind, suggesting regional variations in purchasing power or demand

KEY INSIGHT

- **Product Sales Insight:**
  High-value electronics like refrigerators and laptops generate the most revenue, indicating a consumer preference for durable, big-ticket items.

- **Order Fulfillment Insight:**
  Despite high sales, electronics and home appliances also have the highest cancellation rates, suggesting potential issues in logistics or customer satisfaction.

- **Payment Method Insight:**
  PayPal is the most preferred payment method, while Amazon Pay and Gift Cards are less popular possibly due to user convenience or trust factors.

- **Sales Trend Insight:**
  The sharp rise to $9,540 in sales on February 16, following a low on February 15, highlights the powerful impact of well timed promotions in driving customer purchases.

CONCLUSION

The analysis reveals that high-value electronics such as refrigerators and laptops are key revenue drivers, reflecting strong consumer demand for durable, big-ticket items. PayPal emerged as the most preferred payment method, indicating customer trust in reliable digital wallets. A significant sales spike following a low-revenue day highlights the impact of well-timed promotions in driving purchasing behavior. However, high cancellation rates in top-selling categories suggest potential issues in fulfillment or customer satisfaction that need to be addressed. Additionally, regional sales patterns show that southern and central cities like Miami, Denver, and Houston outperformed coastal cities such as New York, Los Angeles, and San Francisco, indicating distinct geographic differences in consumer purchasing trends..

RECOMMENDATION

1. Prioritize High-Value Electronics

Action: Allocate more inventory/marketing budgets to refrigerators, laptops, and smartphones (top revenue generators).

Rationale: These products dominate sales ($78K, $58K, $48K) and have higher average order values.

2. Improve Order Fulfillment for Electronics & Appliances

Action: Investigate root causes of high cancellation rates (e.g., logistics delays, return policies).

Rationale: Electronics/Home Appliances account for 60%+ of cancellations ($36K), hurting revenue.

3. Optimize Payment Methods

Action: Promote PayPal (24% adoption) and explore incentives for underused methods (e.g., Amazon Pay).

Rationale: Payment preferences impact conversion rates; balancing options can reduce checkout friction.

RECOMMENDATION (PART 2)

4. Leverage Promotional Timing

Action: Replicate the Feb 16 sales spike ($9,540) with targeted campaigns before low-revenue dates.

Rationale: Sharp revenue surges suggest promotions drive demand effectively.

5. Diversify Low-Performing Categories

Action: Bundle low-revenue items (e.g., books, apparel) with high-value products or phase out underperformers.

Rationale: These categories contribute <5% of total revenue ($1K–$2K).

Thank you

Geoffrey Otobor Onajefe

Email: otoborgeoffrey@gmail.com

LinkedIn: www.linkedin.com/in/geoffreyotobor