**ECommerce Data Analysis Final Report**

**Executive Summary**

This report provides a statistical and analytical overview of an 8-year eCommerce dataset comprising transactional and customer behavior insights. The objective is to evaluate revenue performance, category trends, customer dynamics, and temporal sales patterns to guide future strategic decisions.

**1. Revenue & Transactional Performance**

* **Total Revenue (2014–2021):** $105,401,435.75
* **Total Transactions:** 1,000,000 transactions
* **Average Order Value (AOV):** $105.40
* **Number of Stores:** 726 stores across 7 divisions

These figures indicate a robust business scale, with a strong order volume and a healthy AOV across a nationwide presence.

**2. Category-wise Revenue Breakdown**

|  |  |  |
| --- | --- | --- |
| **Category** | **Revenue (USD)** | **Percentage of Total Revenue** |
| Beverage | $36,614,400.75 | 34.74% |
| Food | $32,097,142.50 | 30.45% |
| Coffee Supplies | $16,852,945.50 | 15.99% |
| Kitchen Supplies | $8,434,526.50 | 8.00% |
| Dishware | $8,289,452.50 | 7.86% |
| Gum | $2,448,854.00 | 2.32% |
| Medicine | $664,114.00 | 0.63% |

* **Top-Selling Categories:** Beverage and Food generate 65.19% of the total revenue, confirming their dominance and critical business importance.
* **Low-Performing Categories:** Medicine and Gum contribute less than 3%, signaling either niche roles or underperformance.

**3. Time Series Sales Analysis (2014–2021)**

Sales data exhibits consistent fluctuations, suggestive of seasonal buying behavior or event-based promotions (e.g., holidays, back-to-school seasons).

* **Anomaly in January 2021:** Sales dipped sharply below 900,000 units, breaking prior seasonal patterns.

**Potential Causes:**

* Market Disruption (e.g., pandemic effects, regulatory changes)
* Shifts in consumer preferences
* Operational or supply chain issues
* Competitor activity or pricing strategies

**4. Customer Behavior & Growth Insights**

All customers were added in 2014, with rapid decline month-over-month, implying a closed customer ecosystem, with no new customers after August 2014.

**New Customer Acquisition (2014)**

|  |  |
| --- | --- |
| **Month** | **New Customers** |
| Jan | 3640 |
| Feb | 3857 |
| Mar | 1206 |
| Apr | 354 |
| May | 79 |
| Jun | 39 |
| Jul | 15 |
| Aug | 1 |

**Customer Activity (Last Transactions)**

|  |  |
| --- | --- |
| **Month** | **Unique Customers** |
| Jun 2020 | 1 |
| Jul 2020 | 6 |
| Aug 2020 | 10 |
| Sep 2020 | 55 |
| Oct 2020 | 190 |
| Nov 2020 | 681 |
| Dec 2020 | 2740 |
| Jan 2021 | 5508 |

Despite no new acquisitions post-2014, customer activity surged in late 2020, peaking in January 2021.

**5. Key Insights & Strategic Implications**

* **Customer Growth Concerns:**
  + Auditing data integrity
  + Investing in outreach/marketing
  + Verifying if the dataset is a sample or anonymized scope
* **Revenue Concentration Risk:** Heavy reliance on a few categories can pose risk. Diversification or category innovation might be needed.
* **2021 Sales Drop:** Requires diagnostic investigation—potential causes include market shifts or internal disruption.
* **High AOV:** Suggests that current customers are spending well, but growth must come from expanding the base.

**6. Recommendations**

1. **Data Validation:** Confirm completeness and customer sampling strategy.
2. **Customer Acquisition Strategy:** Relaunch campaigns, loyalty programs, or partnerships to bring in new users.
3. **Category Development:** Analyze underperforming segments for repositioning or promotion.
4. **Anomaly Investigation:** Deep-dive into operational or external factors behind the Jan 2021 drop.
5. **Predictive Modeling:** Use RFM segmentation and clustering to optimize targeting and engagement.