Capstone Project
E-Commerce

Comprehensive Analysis of Sales, Shipping, Customers, and Products

By: Bryan Kim Bauyon

Problem Statement

- Assume that the company is looking to create a customer base, enhancing the customer experience and recommend relevant products to the customers and increase revenue as per the information present.

Research Objective

- To generate a report with some important insights which can be helpful for the company to better understand their customers.

Data & Methodology

- Data Sources:
  - `purchase` table: (orders, prices, discounts, tax, shipping)
  - `customer` table: (demographics, loyalty metrics)
- Methodology:
  - Built KPIs with DAX measures (e.g., Revenue, Delay, Segments).
  - Segmented customers using percentile analysis (High/Med/Low).
  - Visualized results in 7 themed Power BI pages.

Sample DAX Measures used for Analysis in Power BI

Total Revenue	Average Shipping Delay	Customer Segment (Percentile Based)

Total Revenue =
 SUMX (
   purchase,
   purchase[price] * (1 - purchase[discount]) *
   purchase[quantity] * (1 + purchase[tax])
)
	Average Shipping Delay (Days) =
	 AVERAGEX (
	   FILTER (
	     purchase,
	     NOT ISBLANK (purchase[shipping_date]) &&
	     NOT ISBLANK (purchase[order_date])
	   ),
	   DATEDIFF (purchase[order_date], purchase[shipping_date], DAY)
	)
		Customer_Segment =
		 VAR CustRevenue =
		  CALCULATE (
		    [Total Revenue],
		    FILTER (
		     purchase,
		     purchase[customer_id] = EARLIER (purchase[customer_id])
		    )
		  )
		 VAR P80 =
		  PERCENTILEX.INC (
		    SUMMARIZE (purchase, purchase[customer_id], "Revenue", [Total Revenue]),
		    [Revenue], 0.8
		  )
		 VAR P50 =
		  PERCENTILEX.INC (
		    SUMMARIZE (purchase, purchase[customer_id], "Revenue", [Total Revenue]),
		    [Revenue], 0.5
		  )
		 RETURN
		  SWITCH (
		    TRUE(),
		    CustRevenue >= P80, "High Value",
		    CustRevenue >= P50, "Medium Value",
		    "Low Value"
		  )

Dashboard Structure

- **Dashboard Pages (Power BI):**
  1. Executive Overview
  2. Customer Demographics
  3. Regional Performance
  4. Customer Value & Loyalty
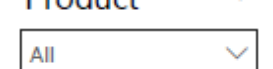 5. Shipping & Operations
  6. Product & Category Performance
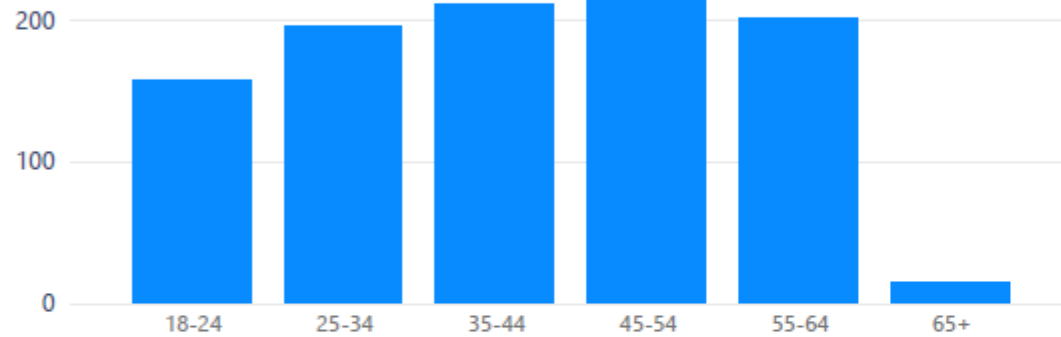 7. Recommendations & Strategy
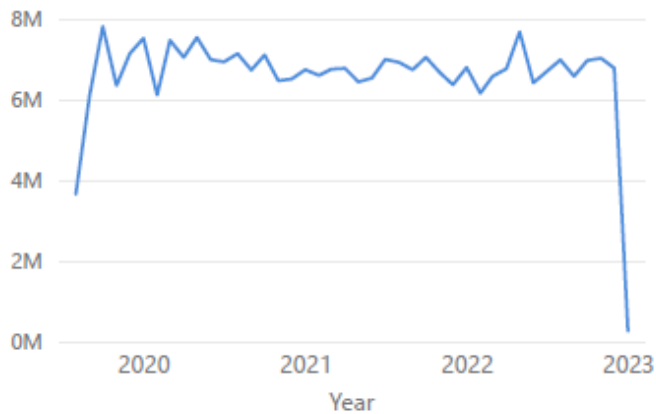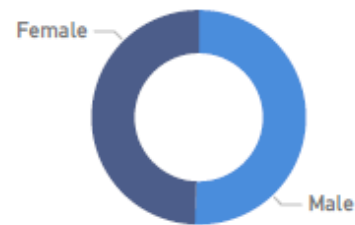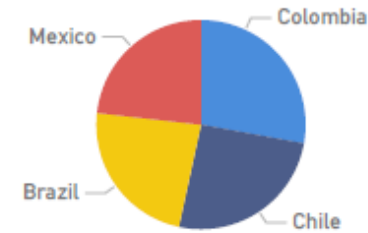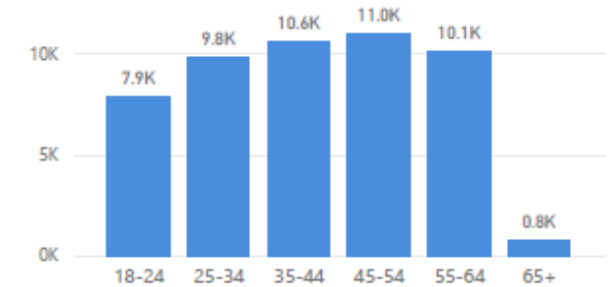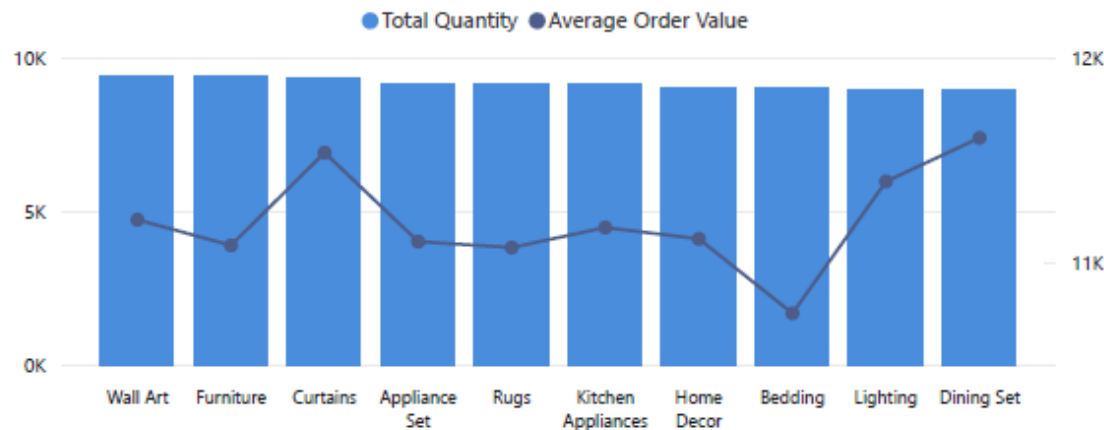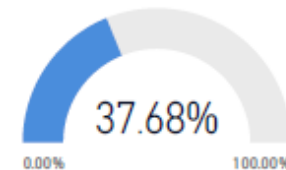
1 Executive Overview

☑ Steady revenue growth across the timeline.

☑ Seasonal spikes highlight peak demand periods.

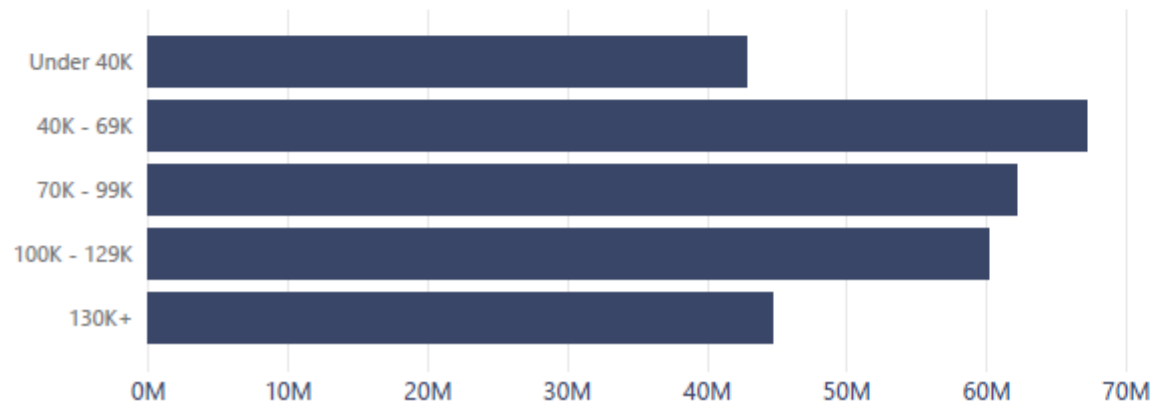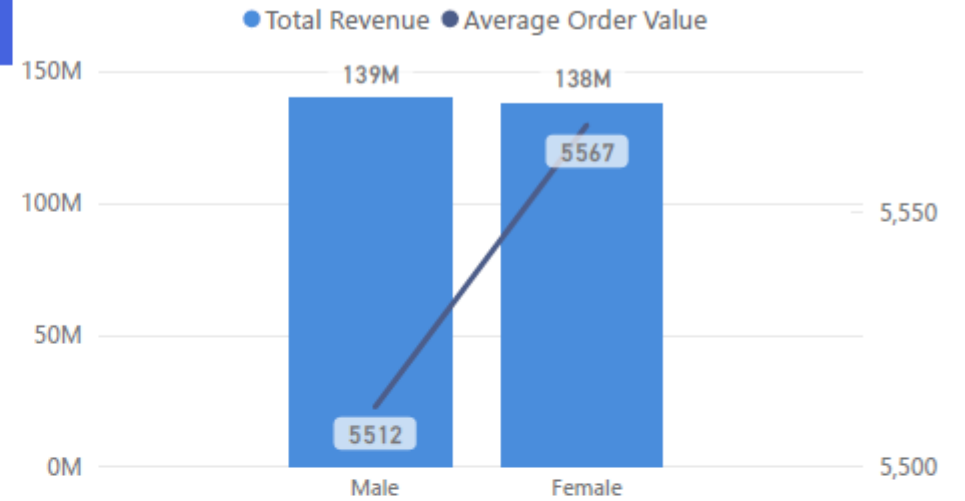⚠ Discounts reduce margins during high-sales months.
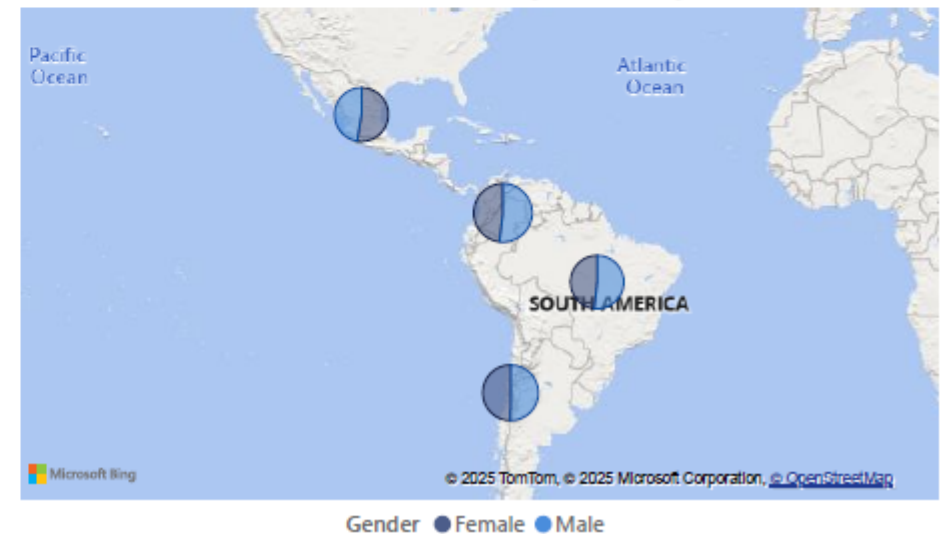
2 Customer Demographics

☑ Majority of customers are in the **25–40 age group**.

☑ Income distribution skewed towards **mid-income brackets**.

⚠ High-value customers cluster in specific segments → need focused loyalty efforts.
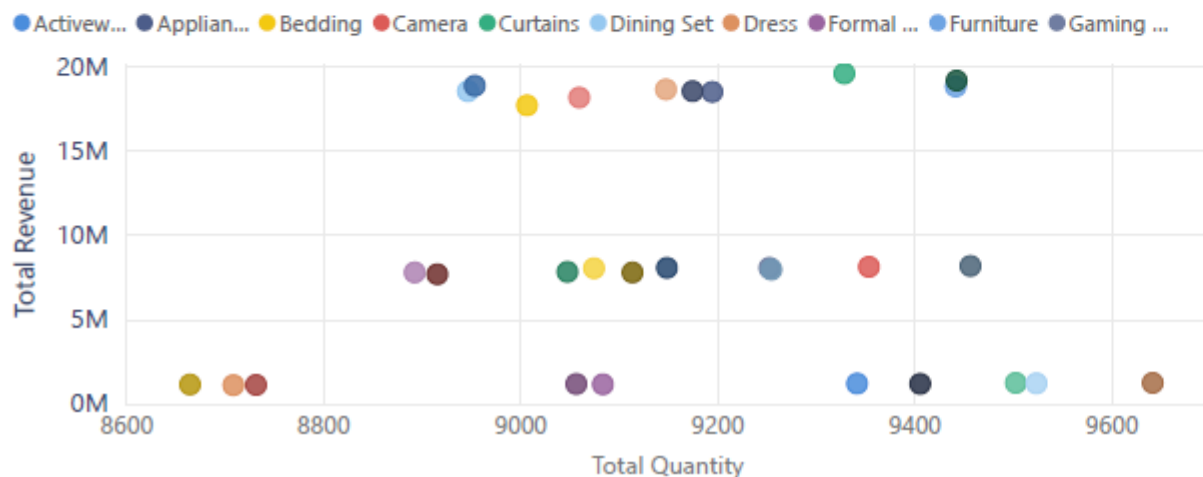
3 Regional Performance

☑ Certain regions (e.g., North/Metro) drive **majority of revenue**.

⚠ Some regions have **high order counts but low revenue** → possible low pricing or weaker product mix.

⚠ Opportunity to expand marketing in **underpenetrated regions**.
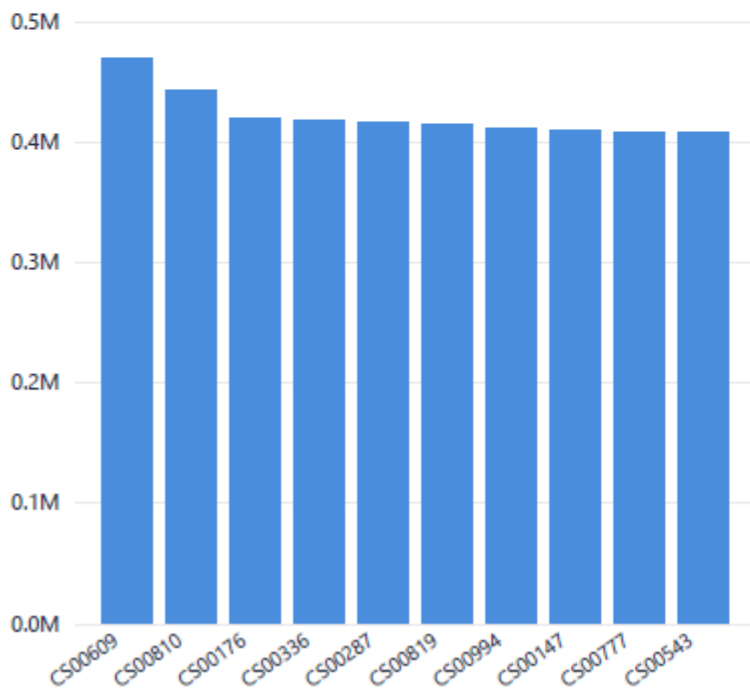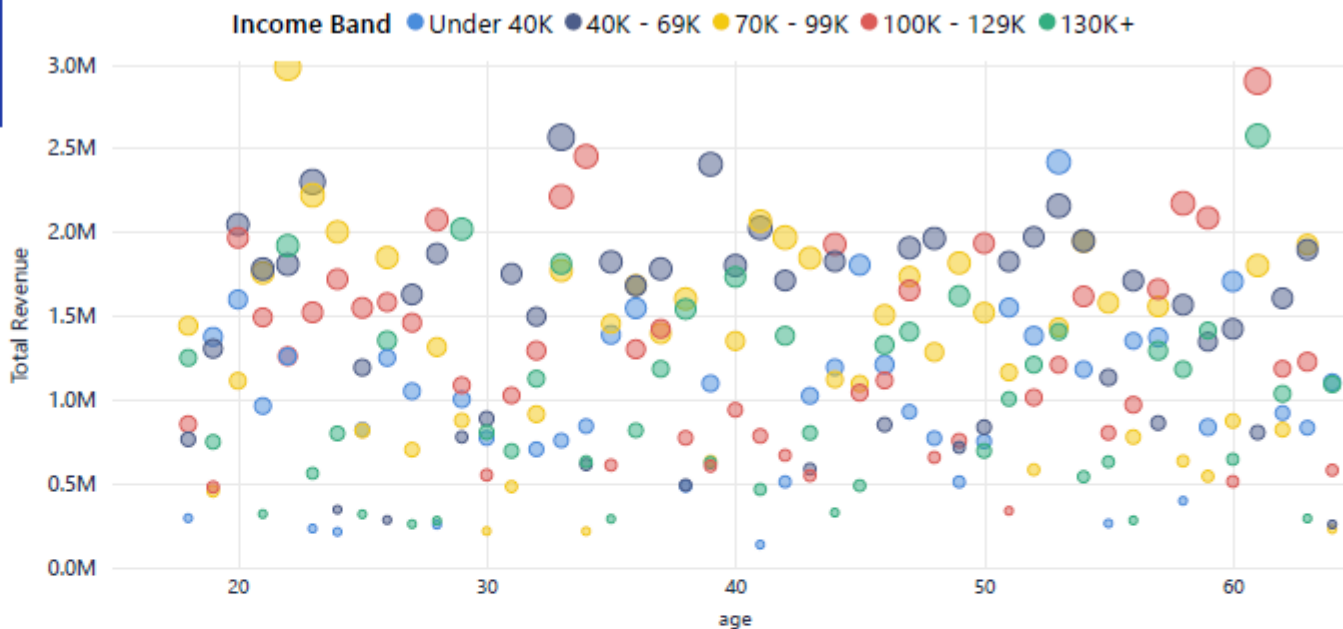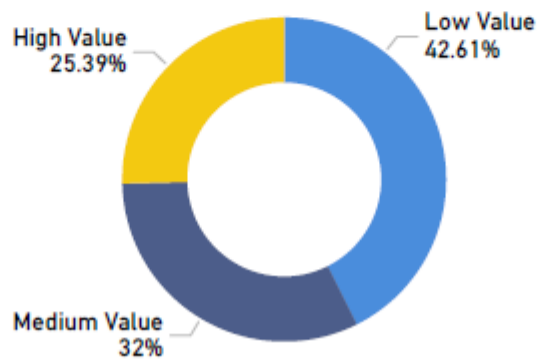
4 Customer Value & Loyalty

☑ Top 20% of customers contribute **~80% of revenue**.

☑ High-value customers (gold segment) show consistent repeat purchases.

⚠ Most customers still fall in **low-value segment** (<50th percentile).

⚠ Loyalty rate lower than expected (few repeat buyers drive most revenue).

5 Shipping & Operations
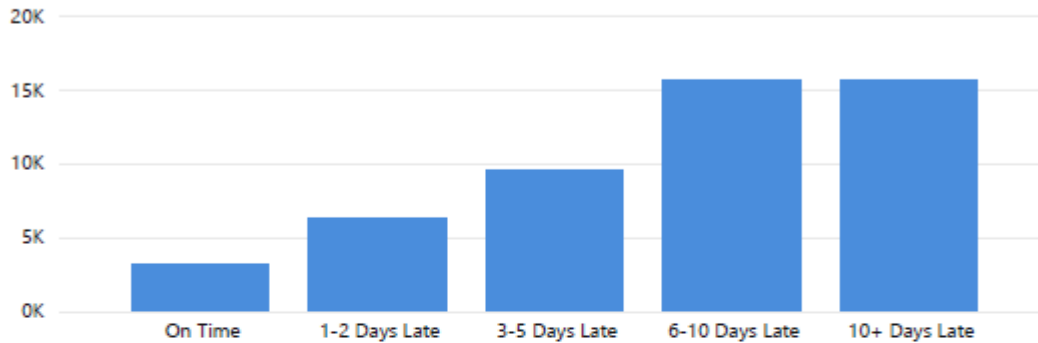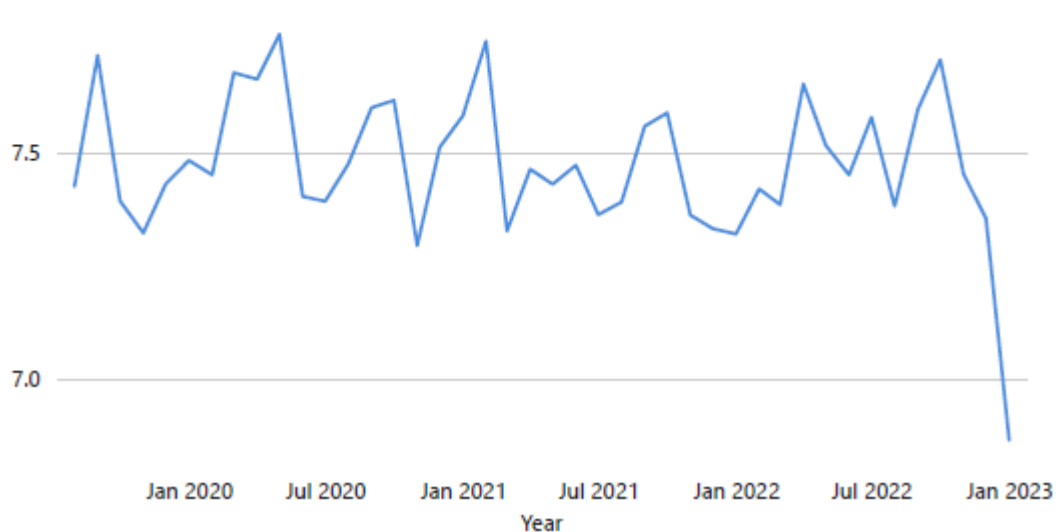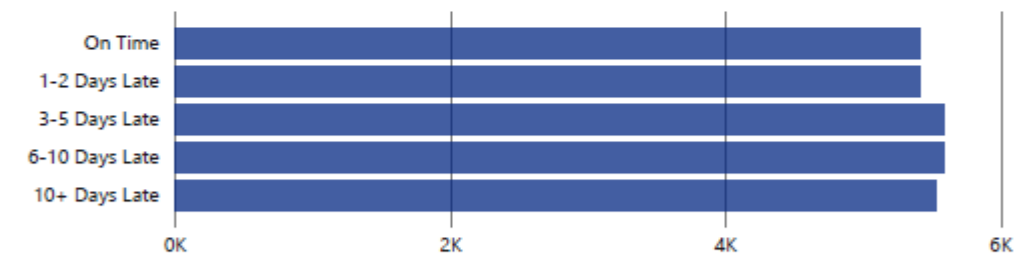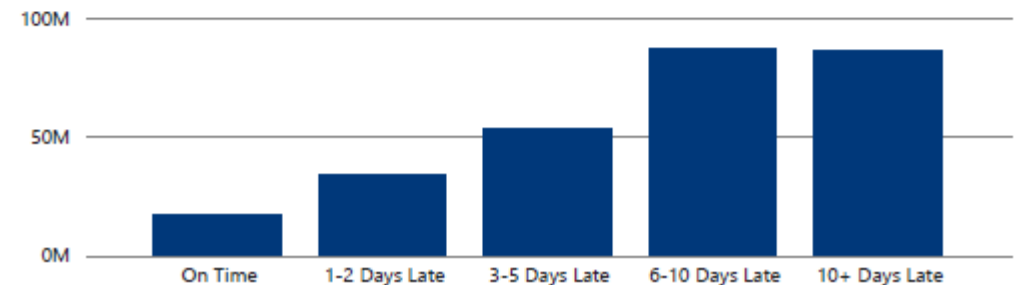
5 Shipping & Operations

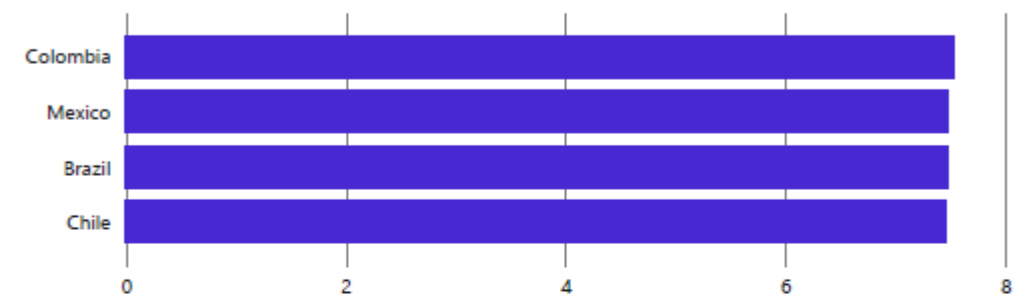☑ Average shipping delay ~**2–3 days**.

☑ On-time delivery rate is **strong (>85%)**.

⚠ Shipping delays linked to **higher costs** in some cases.

⚠ No differentiation by shipping method → harder to optimize operations.

6 Product & Category Performance

☑ Top 3 categories contribute **~70–80% of revenue**.

☑ High demand categories (e.g., Electronics, Home & Furniture).

⚠ Some categories consistently underperform (low revenue & low quantity).

⚠ Product mix heavily skewed → over-reliance on few categories.
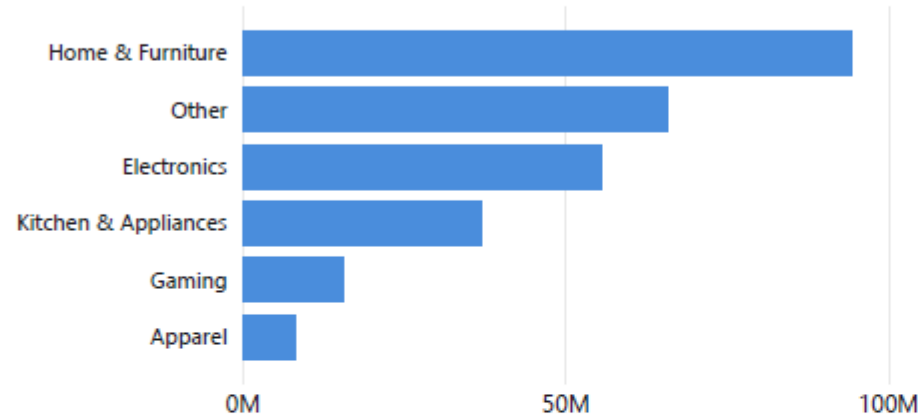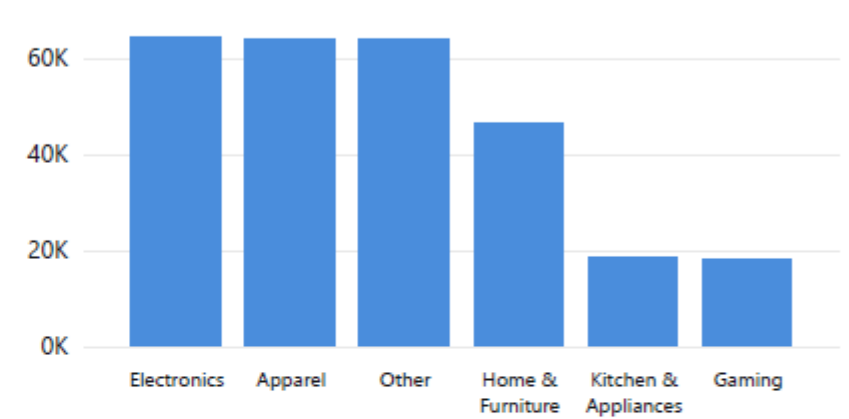
7 Recommendations & Strategy
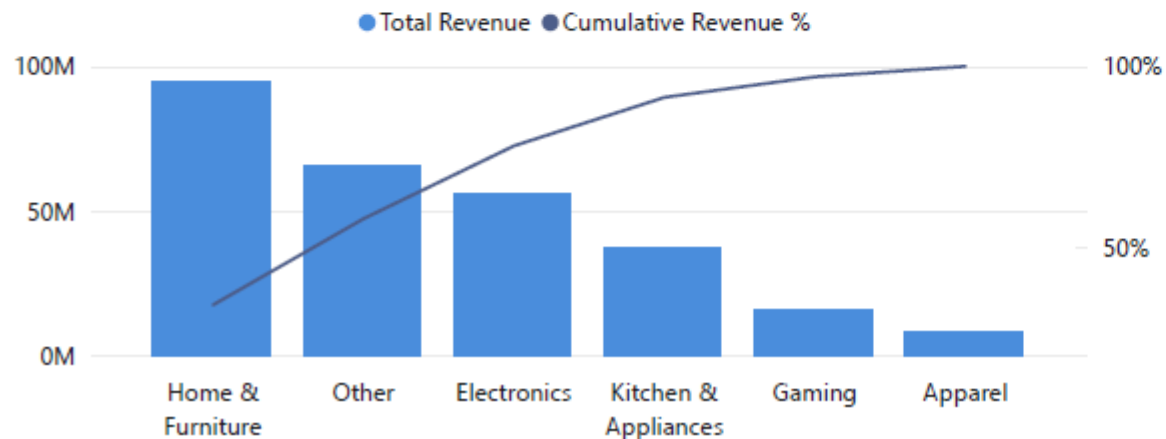
- ☑ **Revenue Focus**: Top 3 categories (Home & Furniture, Other, Electronics) contribute ~78% of revenue → focus marketing & inventory on these.

- ⚠ **Operations Risk**: Ontime delivery rate only ~38%, average delay ~7.5 days → need process improvements & faster shipping.

- ⚠ **Customer Value**: Top 10 customers = ~1.5% of revenue → expand loyalty & retention programs to grow high-value customers.

- ☑ **Growth Opportunity:** Mid-income (40-99K) segments & age bands 25-44 show strong revenue → target promotions here

- ☑ **Diversification**: Underperforming categories (Apparel, Gaming, Kitchen & Appliances) need repositioning, bundle offers, or phase-out.

- ⚠ **Regional Weak Spots:** Address regions with low sales conversion.

Thank you