

# Sales Data Exploratory Analysis

## A Deep Dive into Drivers of Sales Performance

Lucas Okwudishu

2024-06-12

### Table of contents

0.1	Executive Summary . . . . .	2
0.2	Exploratory Data Analysis . . . . .	3
0.3	Sales Trend Deepdive . . . . .	5
0.4	Analysis of Repeat Customers . . . . .	10
0.5	Summary & Recommendations . . . . .	15

### List of Figures

1	Top 10 Sales by Category . . . . .	3
2	Top 20 Sales by Products & Categories . . . . .	4
3	Sales Trend (Online) . . . . .	5
4	Sales Trend (In-Instore) . . . . .	6
5	Sales Trend by Product Category . . . . .	7
6	Sales Trend by State . . . . .	8
7	Sales Trend by Customer Age Groups . . . . .	9
8	Repeat Customers % of Total Customers . . . . .	10
9	Repeat Purchasers by Number of Unique Purchases . . . . .	11
10	Repeat Customers Prop vs Total Sales Trend . . . . .	12
11	Total Customers by Month . . . . .	13
12	Average Sales Value Per Customer by Month . . . . .	14
13	Summary of Customers and Sales by Period . . . . .	15

## 0.1 Executive Summary

This report presents the first phase of a comprehensive analysis aimed at developing marketing and sales strategies for a global electronics retailer, and focuses on online sales in the United States. This phase, titled “**Exploratory Data Analysis**”, examines sales trends from January 2016 to February 2021, segmented into three distinct periods: Pre-Peak Sales (January 2016 - May 2018), Peak Sales (June 2018 - February 2020), and Post-Peak Sales (March 2020 - February 2021).

Key findings from this analysis include:

- **Sales Trends:** Monthly sales peaked in February 2019, with significant spikes in December 2018, May 2019, and August 2019. Post-February 2020, a steady decline in sales is observed.
- **Customer Behavior:** Repeat customers significantly contribute to overall sales, with **90%** of total sales coming from repeat customers. Additionally, **23%** of customers have made at least two purchases and **18%** making three or more purchases.
- **Sales Decline Factors:** The recent sales decline is attributed to a noticeable drop in the total customer count, a slight decrease in the proportion of repeat customers, and a minor decline in average spend per customer.

The analysis concludes that no single customer cohort or product category is driving the sales decline. Instead, multiple factors collectively contribute to the trend. Based on these insights, our recommendation is to implement a comprehensive customer strategy focused on acquiring new customers and encouraging more frequent purchases from existing customers.

---

## 0.2 Exploratory Data Analysis

In this phase, we focus on just understanding aspects of our retail data such as customer segments, products, product categories, and sales trends. To reproduce the analysis, see section 5.0 of this script.

### 0.2.1 Date Range

Our data ranges from January 2016 to February 2020.

### 0.2.2 Top 10 Sales by Category

Top 10 Sales by Product Categories		
Category	Sales (USD)	Sales Prop
Computers	\$2,080,808	34.07%
Home Appliances	\$1,125,427	18.43%
Cell phones	\$779,367	12.76%
Cameras and camcorders	\$722,624	11.83%
TV and Video	\$611,053	10.01%
Audio	\$363,446	5.95%
Music, Movies and Audio Books	\$340,859	5.58%
Games and Toys	\$83,622	1.37%

Figure 1: Top 10 Sales by Category

Computers lead all categories in terms of sales with **\$2M** worth of sales and **34%** of total sales. Except Computers and Home Appliances, all other categories have not hit the **\$1M** mark in terms of sales. Games and Toys on the other end has generated **\$83K** in sales which is only **1.37%** of total sales.

### 0.2.3 Top 20 Sales by Product & Category

**8** of the top 20 products by sales are from the Computers category. The other **2** products are from TV and Video and Home Appliances.

Top 20 Sales by Products & Category			
Product Name	Category	Sales (USD)	Sales Prop
WWI Desktop PC2.33 X2330 Black	Computers	\$85,467	1.40%
Adventure Works Desktop PC2.33 XD233 Brown	Computers	\$75,582	1.24%
Fabrikam Refrigerator 24.7CuFt X9800 Grey	Home Appliances	\$57,600	0.94%
WWI Desktop PC2.33 X2330 White	Computers	\$56,059	0.92%
Adventure Works Desktop PC2.33 XD233 Black	Computers	\$55,233	0.90%
Adventure Works Desktop PC2.30 MD230 White	Computers	\$50,915	0.83%
WWI Desktop PC2.33 X2330 Silver	Computers	\$45,031	0.74%
Adventure Works 52" LCD HDTV X590 White	TV and Video	\$43,500	0.71%
WWI Desktop PC2.33 X2330 Brown	Computers	\$40,436	0.66%
Adventure Works Desktop PC2.33 XD233 White	Computers	\$39,729	0.65%

Figure 2: Top 20 Sales by Products & Categories

## 0.2.4 Top 10 Sales by State

These 10 states account for **51%** of all sales.

Top 10 Sales by States			
State	Sales (USD)	Sales Prop	Sales Prop Cumulative
California	\$623,083	10.20%	10.20%
Texas	\$450,767	7.38%	17.58%
New York	\$385,668	6.31%	23.90%
Florida	\$318,160	5.21%	29.11%
Ohio	\$288,411	4.72%	33.83%
Pennsylvania	\$274,499	4.49%	38.33%
Illinois	\$246,335	4.03%	42.36%
Michigan	\$191,368	3.13%	45.49%
Georgia	\$183,238	3.00%	48.49%
Massachusetts	\$166,000	2.72%	51.21%

## 0.3 Sales Trend Deepdive

### 0.3.1 Sales Trend Overall (Online)

As mentioned earlier, we can see a steady upward trend in online sales with a peak sales period between June 2018 and Feb 2020. Thus, going forward, we can break our analysis timeframe into 3 periods -

- **Pre Peak Sales Period:** January 2016 - May 2018
- **Peak Sales Period:** June 2018 - February 2020
- **Post Peak Sales Period:** March 2020 - February 2021

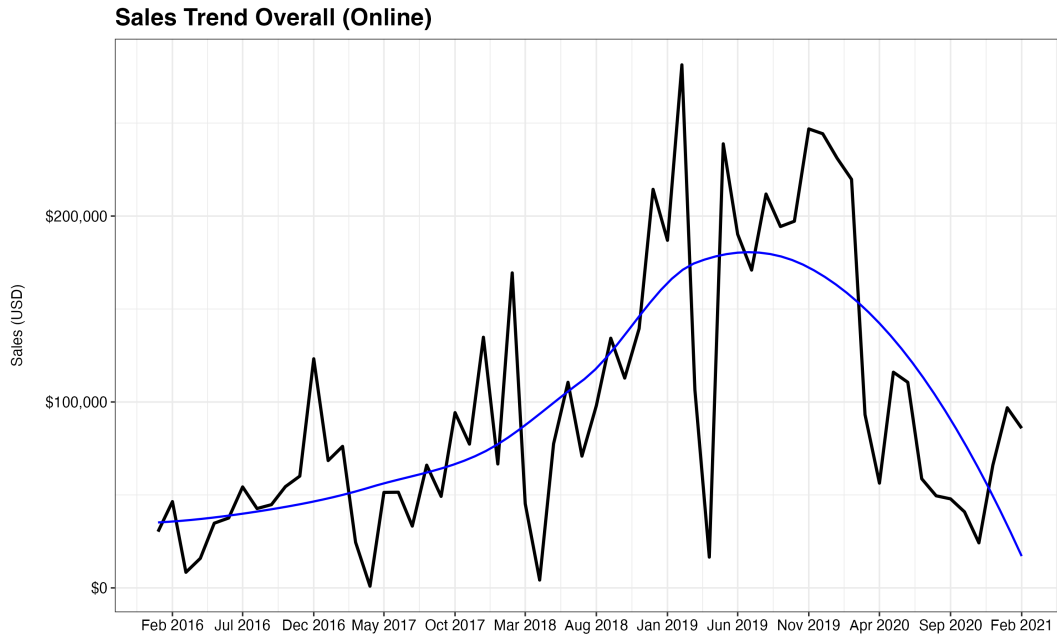


Figure 3: Sales Trend (Online)

Monthly sales peaked in February 2019 with total sales of **\$281K**. Other high points include December 2018 (**\$214K**), May 2019 (**\$238K**) and August 2019 (**\$211K**). However, after Feb 2020 we start to see a steady decline in monthly sales. Going forward, senior management would like to ensure that sales do not dip below pre June 2018 levels.

To investigate this recent sales decline further we can look at sales trends from different angles to see if certain product or customer cohorts may be causing this decline. For example, is there a possibility that customers are starting to prefer in-store sales compared to online? This is highly unlikely. We can expect to see the opposite, customers prefer to shop online and avoid

brick-and-mortar stores. However, to be sure, let's look at the sales trend for brick-and-mortar stores in the US over the same timeframe.

### 0.3.2 Sales Trend Overall (In-Store)

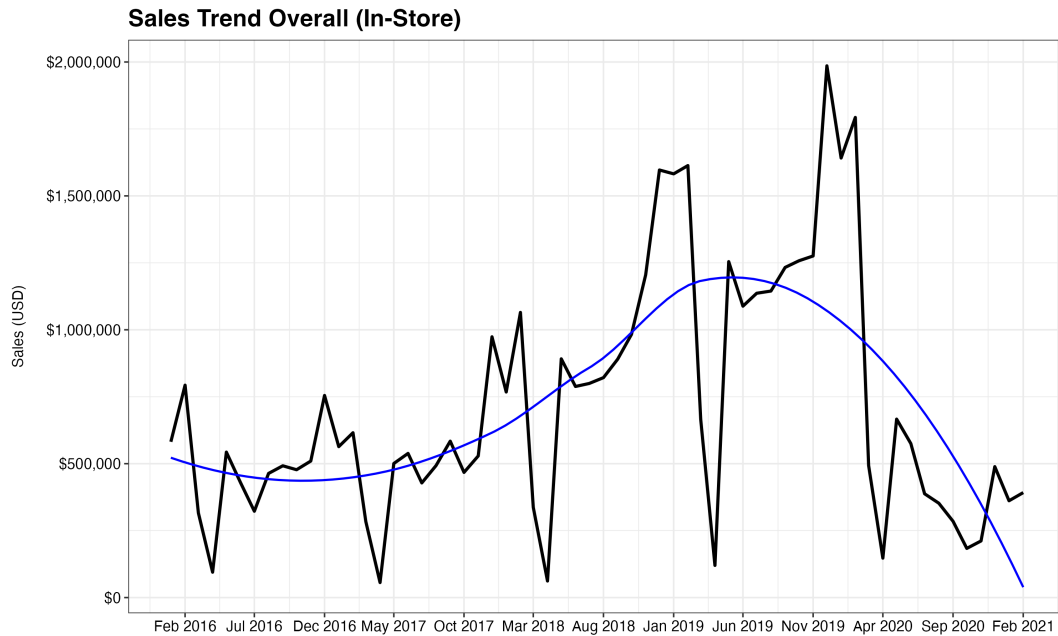


Figure 4: Sales Trend (In-Instore)

From the plot above, we can see the sales trend for in-store sales follows the same pattern as the online sales trend with a peak sales period between June 2018 and Feb 2020 as well. This tells us that in-store sales are currently not cannibalizing the online sales and thus, NOT the reason for the recent sales decline.

### 0.3.3 Sales Trend by Product Category

Are certain product categories responsible for the sales decline?

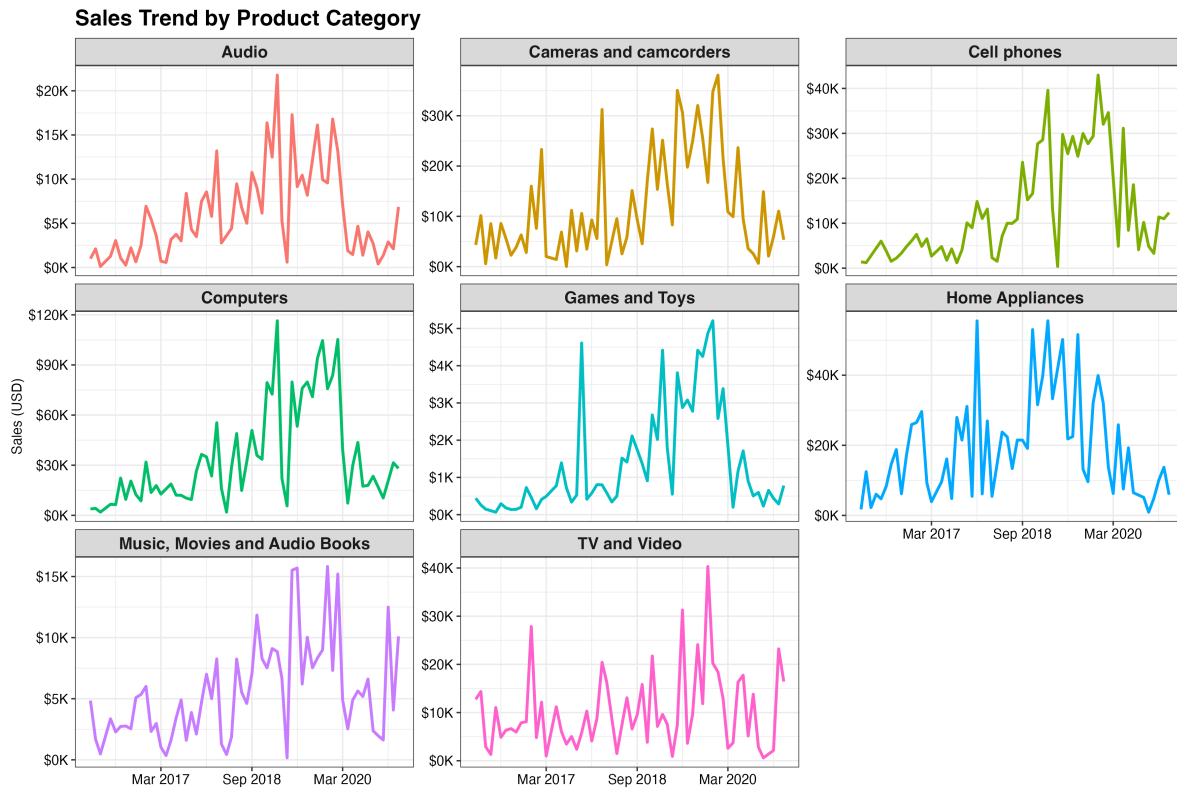


Figure 5: Sales Trend by Product Category

**Unlikely.** From the plot above, we can see that all product categories follow a similar pattern as the overall trend. Some categories like Home Appliances, Game and Toys are declining at a much faster rate than categories like TV and Video, Music, Movies and Audio Books and Cell Phones.

### 0.3.4 Sales Trend by State

Are certain states responsible for the sales decline?



Figure 6: Sales Trend by State

**Unlikely.** Looking at the top 9 states in terms of sales, we also see a similar pattern as the overall trend.



### 0.3.5 Sales Trend by Customer Age Group

Are certain customer age groups responsible for the recent sales decline?

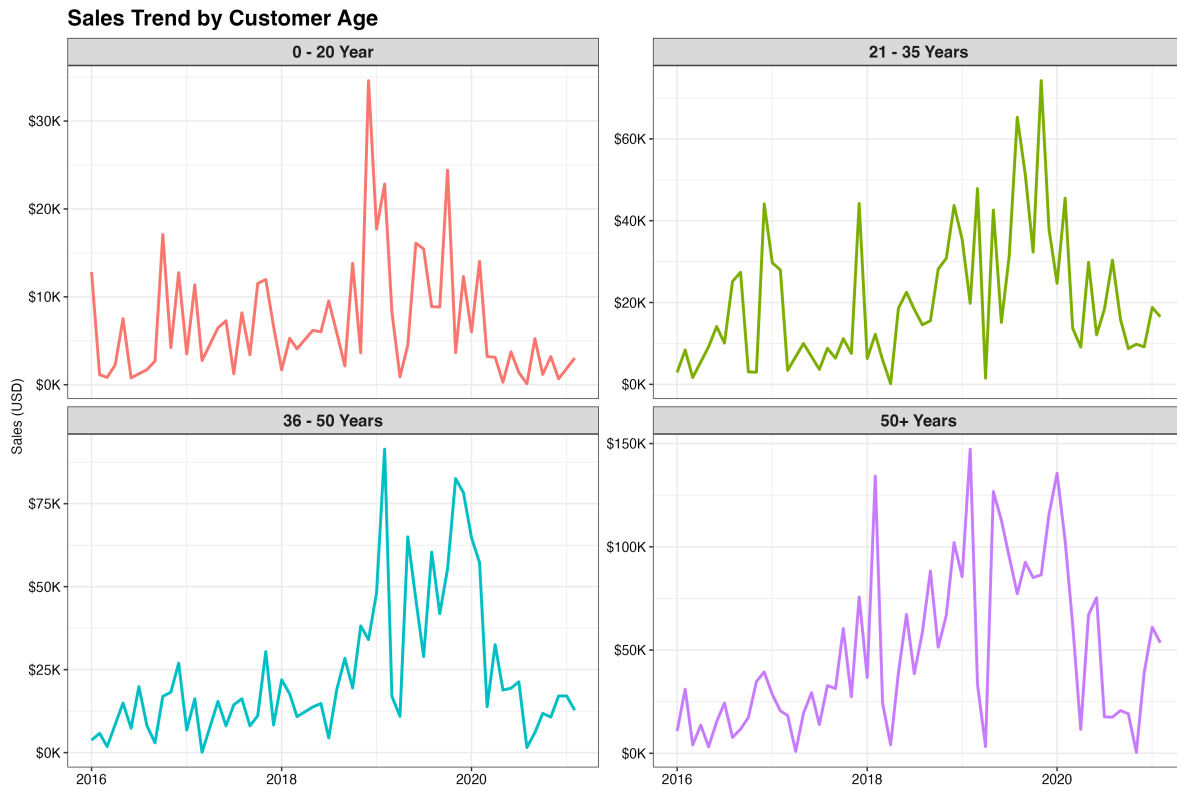


Figure 7: Sales Trend by Customer Age Groups

**Unlikely.** All age groups show a similar pattern as the overall trend. Interestingly the 50+ age cohort appears to have higher sales than all other age cohorts. Normally we would expect the 21 - 35 and 36 - 50 age cohorts to be more tech-savvy age cohorts. In conclusion, the data indicates that no single customer cohort or product category is solely responsible for the recent sales decline.

## 0.4 Analysis of Repeat Customers

Understanding the behavior of repeat purchasers helps in evaluating the health of a retail business. Repeat customers often represent a stable and reliable revenue stream, contributing significantly to sustained sales performance. In this section, we will analyze the impact of repeat purchasers on the overall sales trend, with a specific focus on determining whether the recent sales decline is correlated with changes in the proportion of repeat customers. This will provide insights into the role of repeat customers in the current sales trend and inform strategies for customer retention and sales growth.

### 0.4.1 Sales vs Proportion of Repeat Customers (Overall)

Proportion of Customers with Repeat Purchases					
Total Customers	Total Sales	Repeat Customers	Repeat Customers Total Sales	Repeat Customers Prop	Repeat Customers Sales Prop
2409	\$6,107,205	1761	\$5,605,666	73.10%	91.79%

Figure 8: Repeat Customers % of Total Customers

With 73% of customers making repeat purchases and also accounting for a staggering 91% of total sales, it is evident that loyal customers are the backbone of the retailer's revenue stream. Next, we can also look at the distribution of repeat purchasers by number of purchases.

Next, we can also look at the distribution of repeat purchasers by number of purchases.

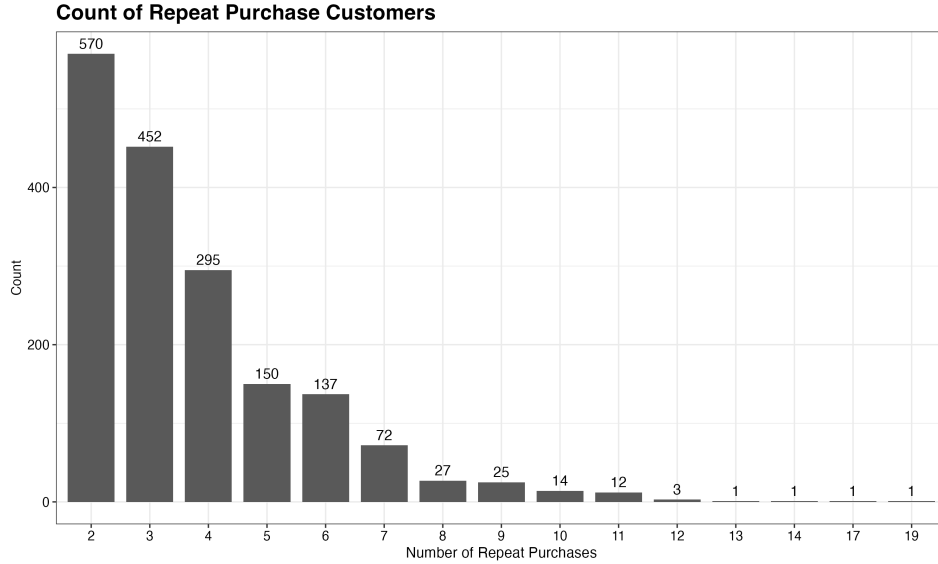


Figure 9: Repeat Purchasers by Number of Unique Purchases

**23%** of customers have made at least two purchases and **18%** have made three purchases. This indicates that repeat customers significantly contribute to the business's success. The large proportion of sales from repeat purchasers suggests that customer retention is not a major issue.

However, the recent decline in sales raises important questions. Since the majority of sales are driven by repeat customers, a decrease in these loyal customers could significantly impact overall sales. This highlights the need to improve strategies not only to retain existing customers but also to encourage them to make more frequent purchases, helping to prevent potential drops in overall sales.

#### 0.4.2 Sales vs Total Customers Trend (Monthly)

Throughout the observed period, the proportion of repeat customers remained steady, often exceeding 60% and peaking around 75%. Despite the normalization of sales after March 2020, the proportion of repeat customers remained substantial. However, the decline in total sales post-March 2020 coincides with a slight decrease in repeat customers, reaching its lowest point of 60% in November 2020.

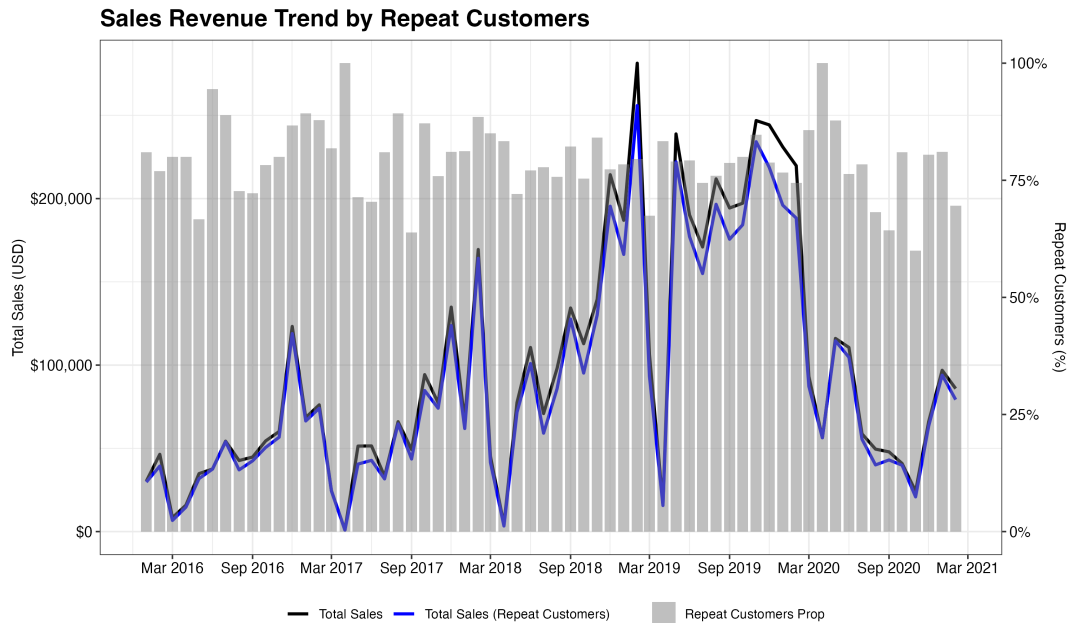


Figure 10: Repeat Customers Prop vs Total Sales Trend

If we look at our 3 key sales periods, Pre-Peak Sales (before June 2018), Peak Sales (June 2018 - February 2020), and Post-Peak Sales (after February 2020), we see that the average proportion of repeat customers was 77.5% during the Post-Peak period, compared to 78.3% during the Peak period and 80% during the Pre-Peak period. During the Pre-Peak period, the high proportion of repeat customers likely did not directly result in increased sales, as this was a ramp-up period for acquiring new customers and familiarizing them with our products. However, in the Post-Peak period, we would expect the high proportion of repeat customers to positively impact sales, but that does not seem to be the case here.

### 0.4.3 Total Customers Trend (Monthly)

The plot below shows the total customers monthly.

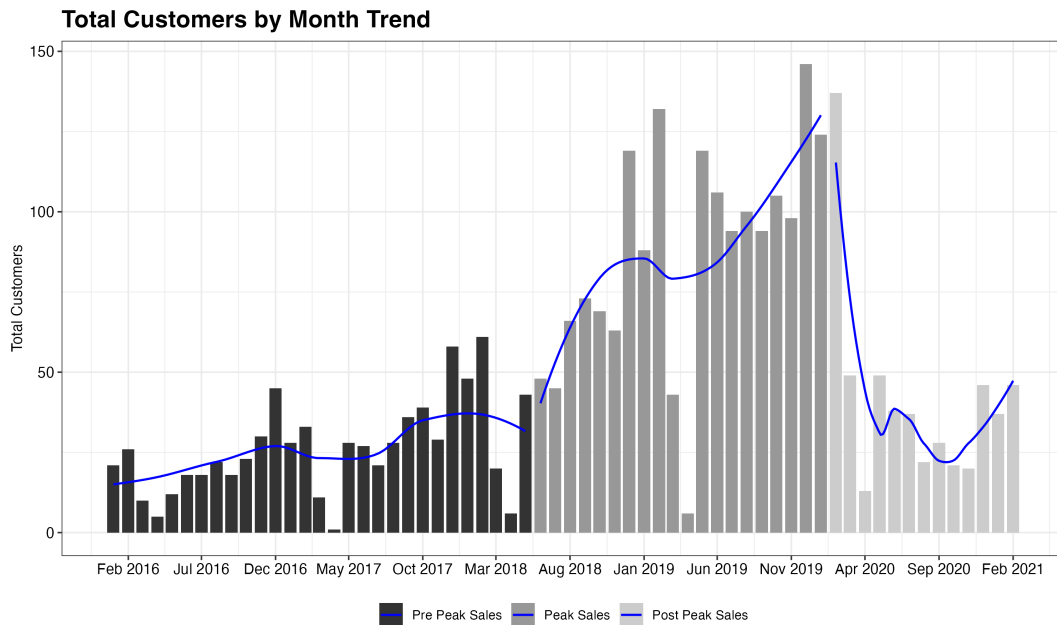


Figure 11: Total Customers by Month

During the Pre-Peak Sales Period, the number of customers per month shows a steady upward trend, despite seasonal fluctuations. Starting from a relatively low base, there is a gradual increase in customer count, reflecting the initial growth phase of the retailer. By May 2018, the customer base had grown significantly, setting the stage for the subsequent peak period.

The Peak Sales Period is characterized by a substantial surge in the number of customers. This phase represents the retailer's period of exponential growth, driven by successful marketing strategies, product launches, or external factors.

The Post-Peak Sales Period shows a noticeable decline in the number of customers per month. By early 2021, customer numbers have not returned to the peak levels observed in 2019, highlighting the challenge of sustaining growth post-peak. The declining trend suggests a need to refocus efforts on acquiring new customers, to regain momentum and stabilize customer counts.

#### 0.4.4 Average Sales per Customer Trend (Monthly)

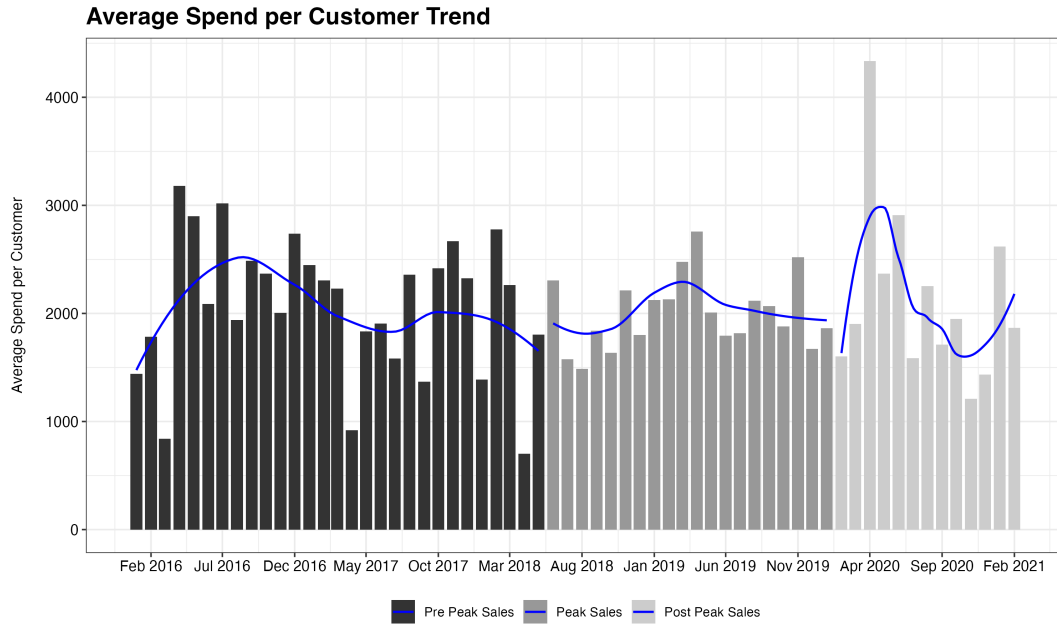


Figure 12: Average Sales Value Per Customer by Month

The analysis of customer behavior from February 2016 to February 2021 reveals distinct phases in spending and customer trends. During the Pre-Peak and Peak Sales Periods, the retailer experienced steady growth and high average spending per customer, with the occasional seasonal fluctuation. However, the Post-Peak Sales Period shows a decline in customer numbers and increased variability in average spending, suggesting challenges in maintaining customer engagement and spending levels. These insights highlight the critical need for targeted strategies to re-engage customers and stabilize spending, ensuring sustained revenue growth and mitigating the impact of the recent sales decline.

## 0.5 Summary & Recommendations

The analysis of customer behavior and spending reveals important insights into the sales trends of our global electronics retailer. The data shows distinct phases: steady growth in the Pre-Peak Sales Period, significant spikes in the Peak Sales Period, and normalization in the Post-Peak Sales Period. However, without additional context on sales strategies, promotions, and external market conditions during the Peak Sales Period, we cannot conclusively determine the causes behind the increased sales observed during this phase. Based on the data, our hypothesis attributes the recent sales decline to several factors.

Summary Total Sales & Customers by Period			
Period	Total Customers	Avg Repeat Customers Prop	Avg Sales Per Customer
Pre Peak Sales	617	80.93%	\$2,664
Peak Sales	1362	78.29%	\$2,495
Post Peak Sales	406	75.59%	\$2,487
Note: Data excludes April 2020 which appears to be an outlier month.			

Figure 13: Summary of Customers and Sales by Period

1. There is a noticeable decline in the total customer count during the Post-Peak Sales Period.
2. Secondly, there is a slight decline in the proportion of repeat customers, which has traditionally been a key driver of sales.
3. Lastly, there is a minor decline in the average spend per customer in the Post-Peak Sales Period, which contributes to the overall sales decline but may not be as significant as the first two factors.

To address our initial goal of ensuring that sales do not dip below Pre-Peak Sales Period levels, our recommendation is to implement an overall customer strategy focused on both acquiring new customers and encouraging existing customers to purchase more frequently. This strategy should involve targeted marketing campaigns, loyalty programs, and personalized engagement to re-engage past customers and attract new ones, ensuring sustained sales growth and stability.