EDA Report for Analyzing Customer Behavior based on Transaction Data

Analyze customer behavior based on transaction data is essential for understanding the purchasing patterns and preferences of the customer base. This EDA report aims to unveil insights from the transaction data to optimize marketing strategies and enhance customer experience.

By NIKHIL JAIN

Markets and Transactions

Business Understanding and Overview

1   **Identify Business Goals**

The goal is to perform Exploratory Data Analysis (EDA) to derive valuable insights into customer into customer preferences, purchasing patterns, and overall behavior..
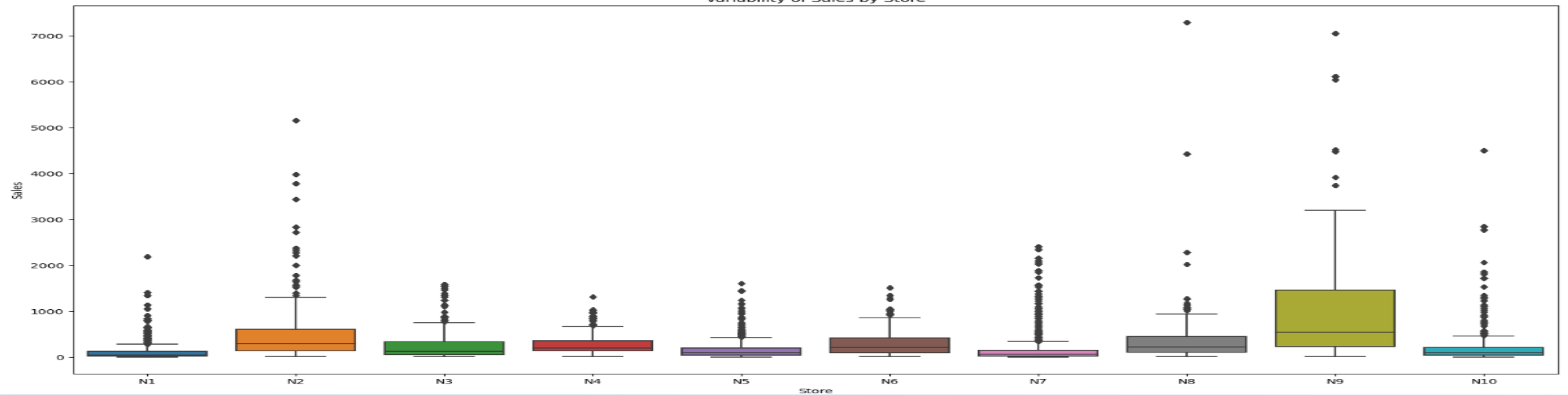
2   **Data Collection**

https://www.kaggle.com/datasets/iamprateek/store-transaction-data

3   **Overview of Market Landscape**

Develop a data analysis project that focuses on understanding and analyzing customer behavior based on transactional data

Variability of Sales By Store

Understanding the Transaction Data

1

The dataset contains information regarding the sales of ten different stores during a period of three months. Our objective is to discover the particular characteristics of each store's sales and possibly determine strategies to implement based on our findings.

2

In the dataset we can see that each transaction has its own Bill_Id number. However, some of them are used by more than one store. In order to avoid miscalculations we'll create a unique Id by combining the Bill_Id with the Storecode number in a new column in order to calculate sales accurately..

Initial Intuition from Data

1

There is no missing values.

2

There are total 14 columns initially .

numerical and nine categorical columns

3

In the dataset we can see that each transaction has its own Bill_Id number. However, some of them are used by more than one store. In order to avoid miscalculations we'll create a unique Id by combining the Bill_Id with the Storecode number in a new column in order to calculate sales accurately.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 26985 entries, 0 to 26984
Data columns (total 14 columns):
 #   Column     Non-Null Count   Dtype
---  ------     --------------   -----
 0   MONTH      26985 non-null   object
 1   STORECODE  26985 non-null   object
 2   DAY        26985 non-null   int64
 3   BILL_ID    26985 non-null   object
 4   BILL_AMT   26985 non-null   float64
 5   QTY        26985 non-null   float64
 6   VALUE      26985 non-null   float64
 7   PRICE      26985 non-null   float64
 8   GRP        26985 non-null   object
 9   SGRP       26985 non-null   object
 10  SSGRP      26985 non-null   object
 11  CMP        26985 non-null   object
 12  MBRD       26985 non-null   object
 13  BRD        26985 non-null   object
dtypes: float64(4), int64(1), object(9)
memory usage: 2.9+ MB
```

Sales by Store Analysis

Findings

- We can observe that stores 7, 9, 5, 2, and 4 are well above the rest in terms of total sales during the three month period.

Total Sales by Store

Quantity of Items Sold by Store
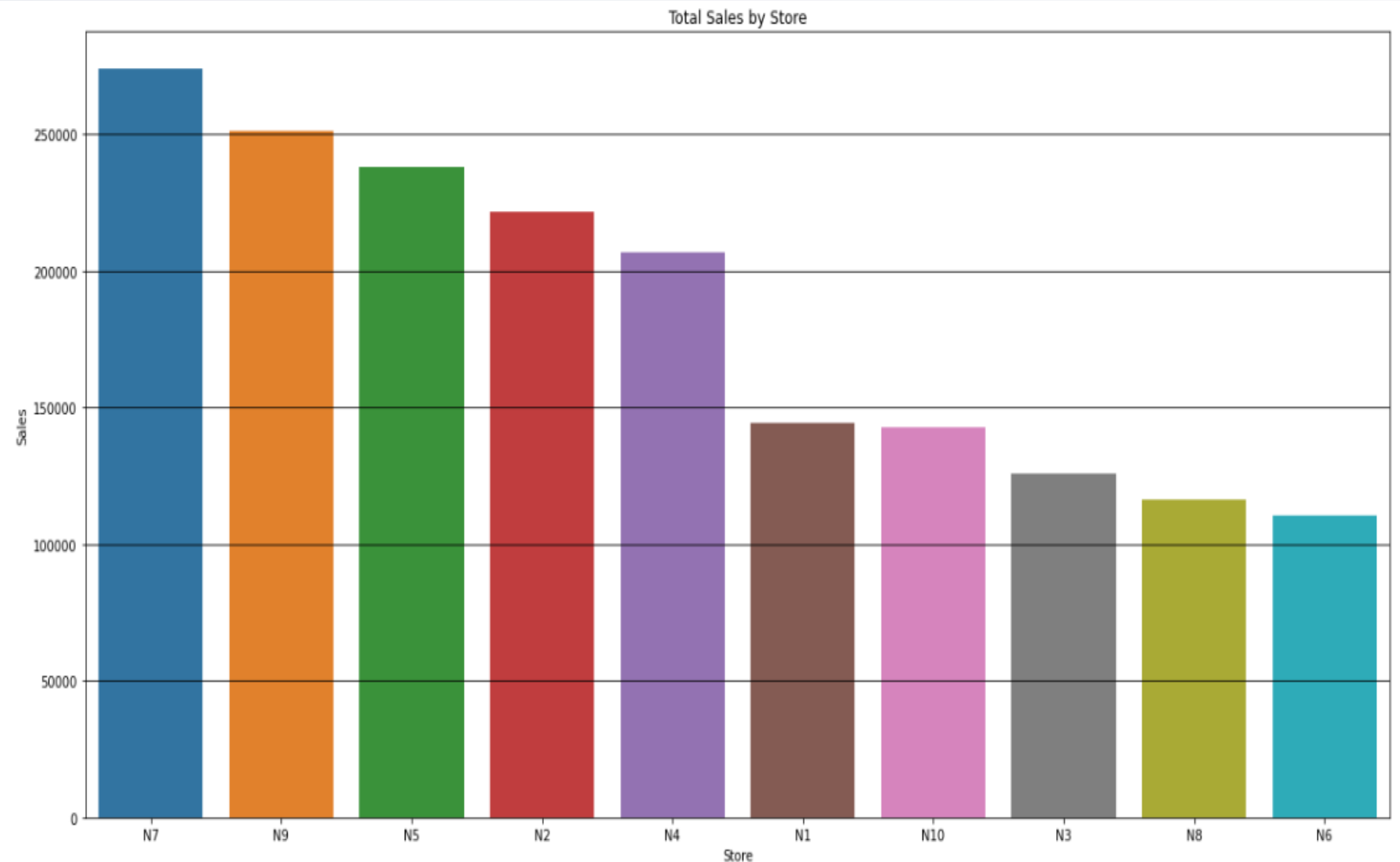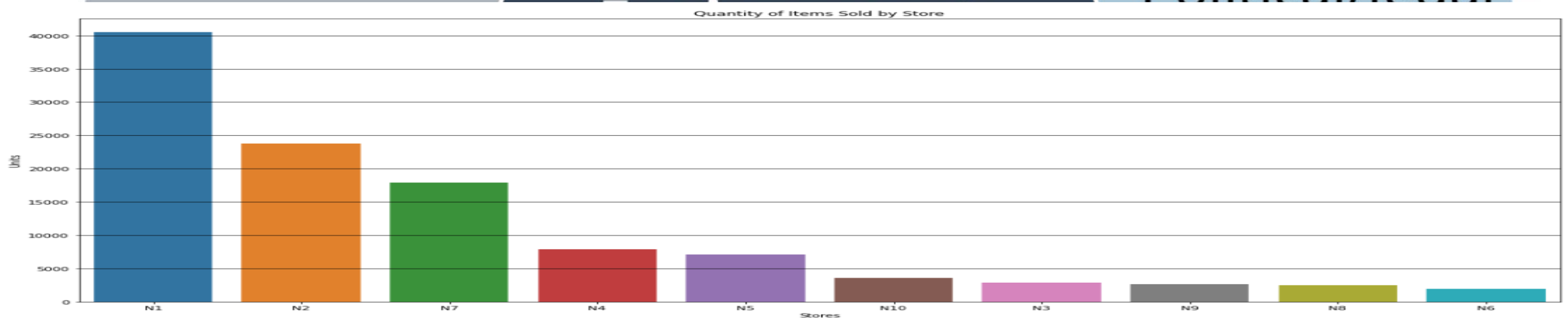
Which Store Sells More Units?

1   **Finding**

This graph shows that store 1 handles five times more volume than most stores. However, as our sales by unit analysis has shown us, store 1's sales by unit is relies largeluy in the sales of biscuits. Therefore, despite handling a seemingly large volume, store 1 may not necessarily be the bigger one.

2   **Finding**

The second biggest store bby units sold is store 2 whose total sales are dominated by packaged tea and coffee. However, their sales by unit mostly rely on pantry products and cleaning products. we can then deduce than it is very likely that store number 2 is indeed a physically larger store.

3   **Finding**

In terms of store 7 (third largest by volume), most of its sales rely in pantry products and personal care products. This store is also the one with the highest sales during the three months analyzed. Stores 4 and 5 have very diversified sales. However, their sales by units have biscuits as a very large component which may help explain why they handle almost twice the unit volume than stores 10, 3, 9, 8 and 6.

Sales by Month Analysis

Findings

- Its sales in month 2 were the highest among the other months, including sales at all stores.
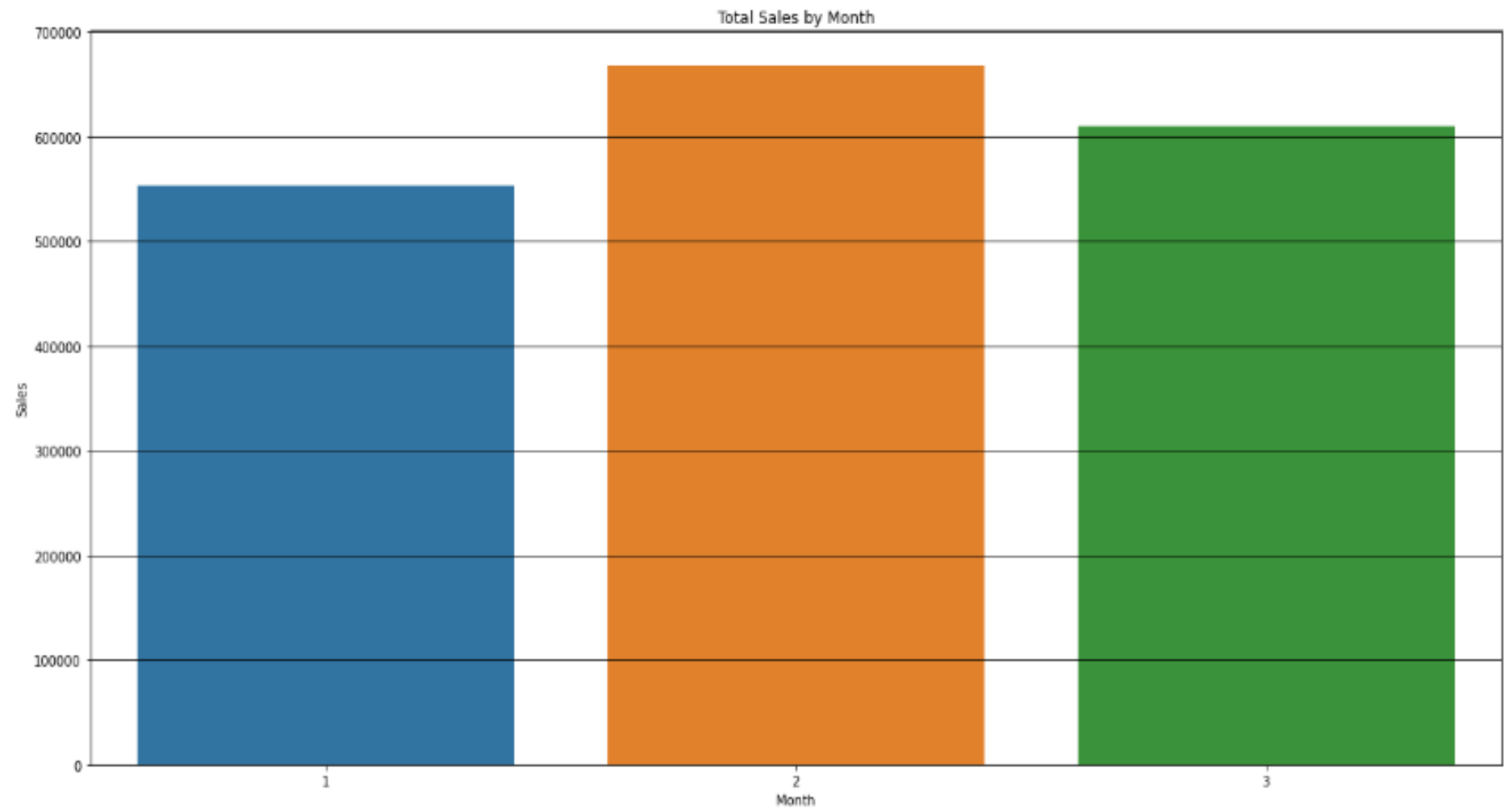
Final Thoughts:

Insights Consolidation

1 Stores 7, 9, 5, 2, and 4 consistently outperform others in terms of total sales.

2 Sales in Month 2 were the highest among all other months across all stores.

3 Suggests a potential seasonal trend or specific events driving increased consumer spending during Month 2.

Recommendations Development

1 Consider store size, product mix, and customer preferences when planning promotions and marketing campaigns.

2 Optimize product placements in larger stores to enhance customer engagement and maximize sales.

3 Develop tailored strategies for each top-performing store based on its unique characteristics and sales patterns.