Academic Report
Machine Learning (CS539)

EDA

Ehu Shubham Kishore Shaw

Prof Yanhua Li's

September 23, 2024

1. **Introduction**

   This research examines transaction and purchase activity data to learn more about product preferences, sales trends, and consumer categories. The data was created by merging client demographics and transaction records to produce an extensive dataset. To investigate different facets of sales success, a number of visualizations were created.

2. **Data Preparation and Cleaning with Data Merging and Feature Engineering**

   - To guarantee a balanced distribution throughout the dataset, missing values in the LIFESTAGE and PREMIUM_CUSTOMER columns were filled during the data preparation stage using predefined category lists. The LYLTY_CARD_NBR was used to combine transaction and customer behavior data from the databases. Furthermore, the DATE column—which was originally presented in an unintelligible minute format—was converted into a standard datetime format, which made precise time-based analysis possible in the visualizations that followed.
   - Few Feature Engineering includes adding column like BRAND and PACK_SIZE to help my analysis allowing me create better results. And visualize it.

3. **Visualization and Analysis**

   1. Customer Segment Analysis.
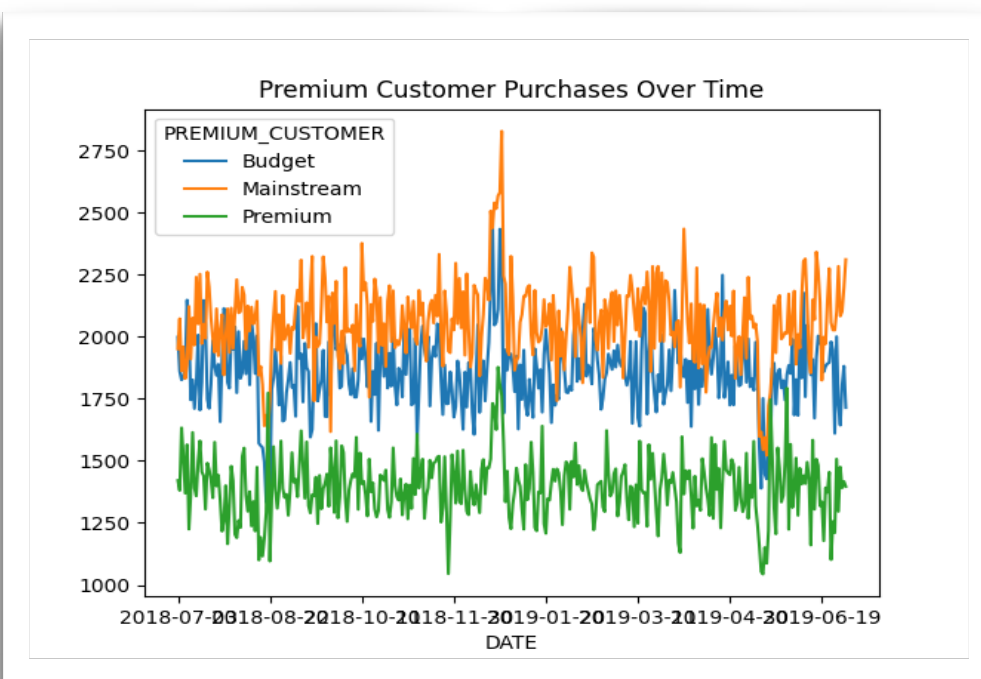
•**Explanation**: This chart shows the proportion of total sales contributed by different customer segments, particularly premium vs. non-premium customers. It highlights which customer groups spend more, helping to identify the most profitable segments based on their purchase.

2. Product Level Analysis .


Total Sales by Packet Size
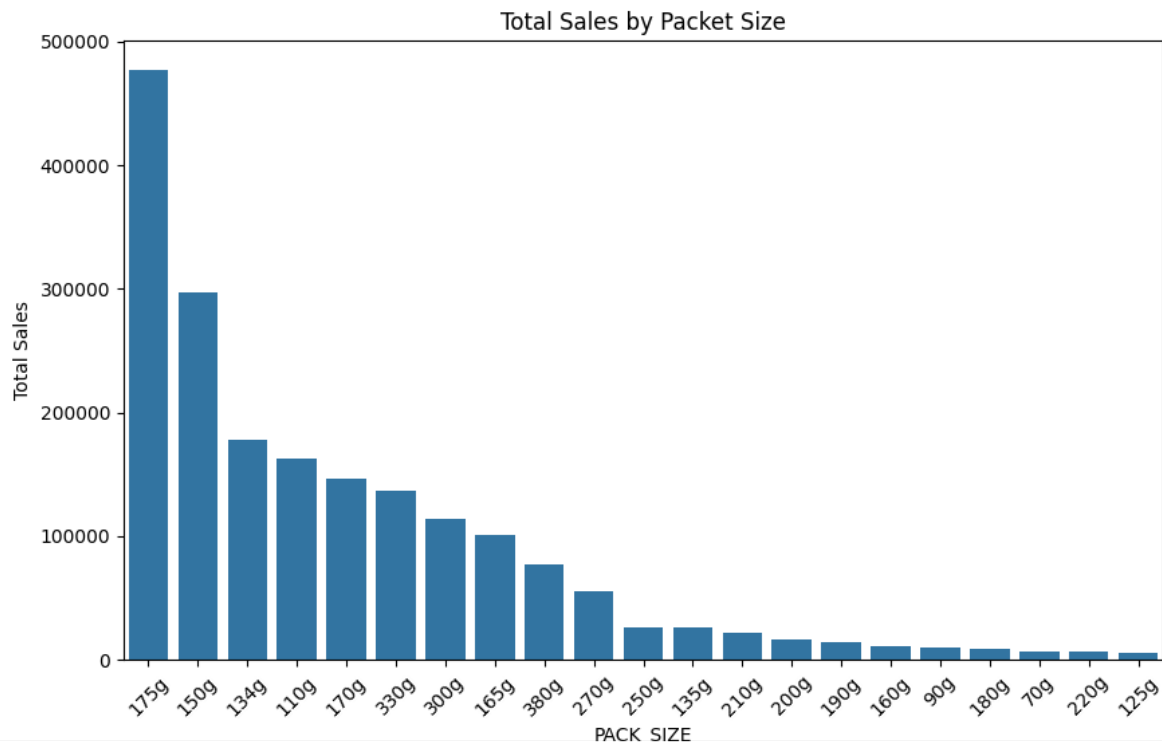
- **Explanation**: This plot displays the total sales based on product PACK_SIZE. It reveals customer preferences for certain pack sizes, showing that larger or specific sizes contribute more to sales, indicating possible bulk-buying behavior.

## 3. Sales Trends Over Time.

Total Sales Over Time
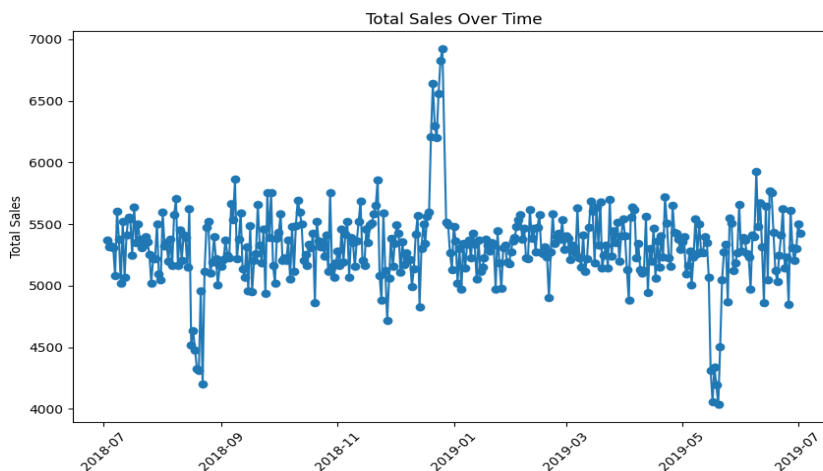
- **Explanation**: This line chart tracks total sales over time, illustrating trends such as seasonal spikes or dips. It helps identify periods of high or low sales activity, which can be linked to events, holidays, or promotions.

4. Sales by Store.
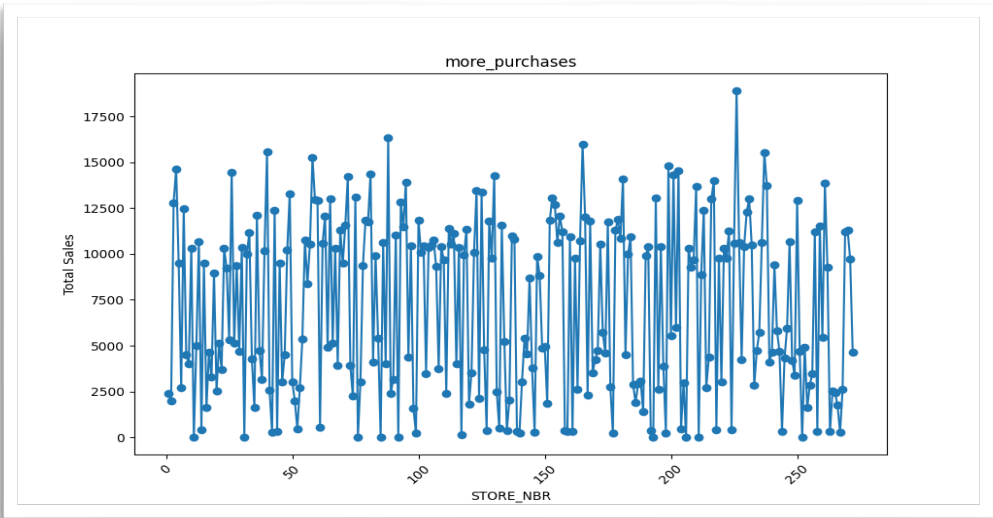
•**Explanation**: This visualization shows total sales per store, helping to identify which stores perform best. Stores with higher sales may be in prime locations or attract more customers, helping in decision-making for store management.

5. Brand Sales Analysis.
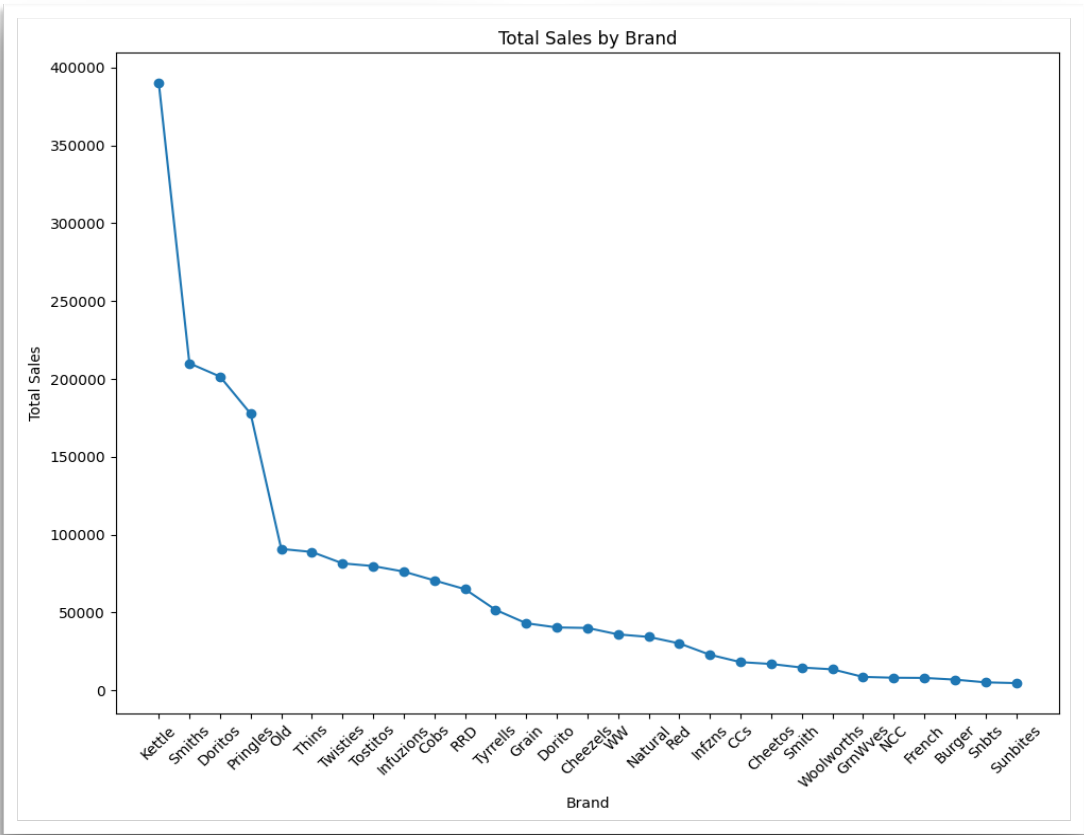
• **Explanation**: This bar plot displays the total sales generated by different brands. It shows which brands are most popular among customers, helping businesses stock more of the high-demand products to maximize revenue.
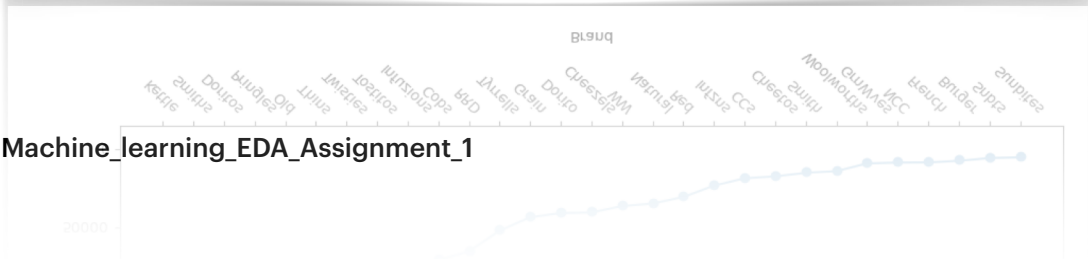
6. Premium Customer Analysis .

- **Explanation**: This chart illustrates the purchasing patterns of premium customers by brand. It highlights which brands are favored by premium customers, providing insights into targeting strategies for higher-value customers.
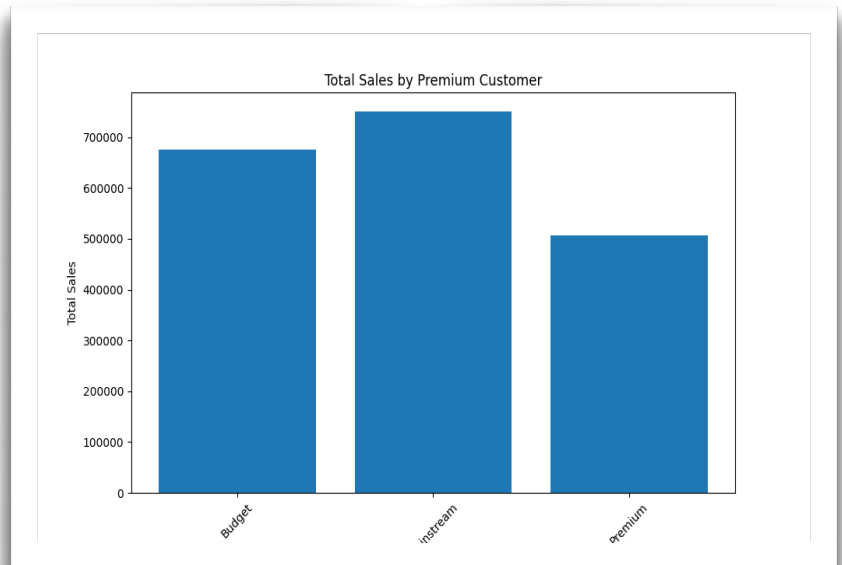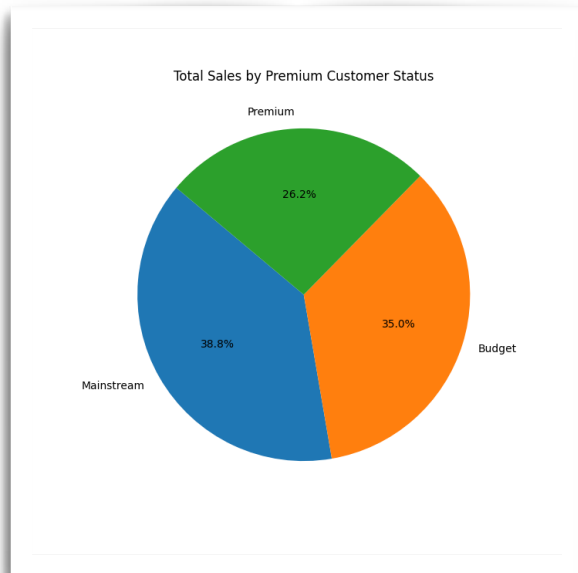
4 . Customer Segment Analysis

- **T**he pie chart shows premium customers contribute only 26% of total sales, indicating lower purchasing activity. The budget segment is balanced, showing mid-level purchasing. Focusing more on premium customers could help boost
overall sales, based on this analysis.
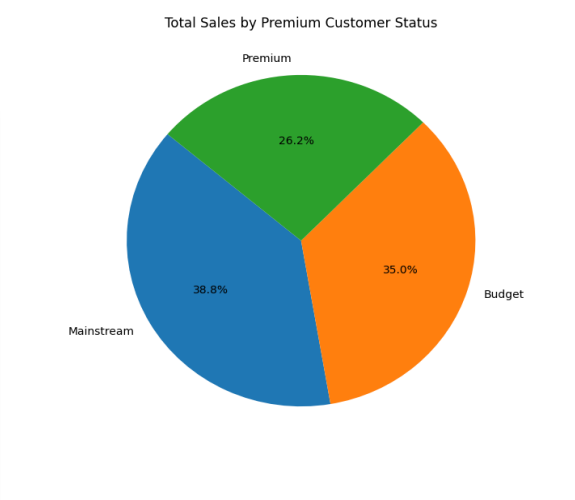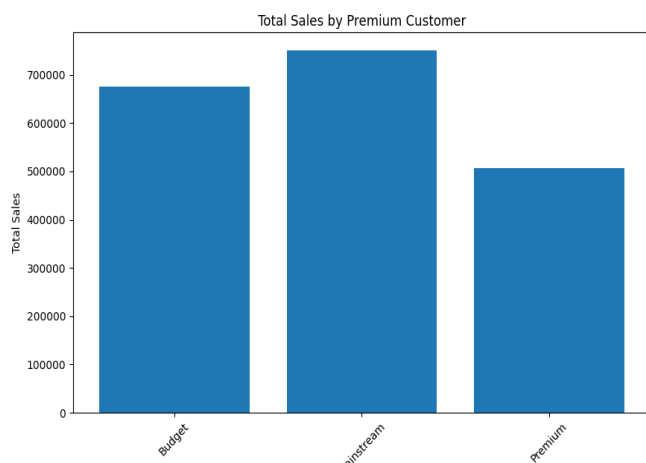
5. Product Level Analysis

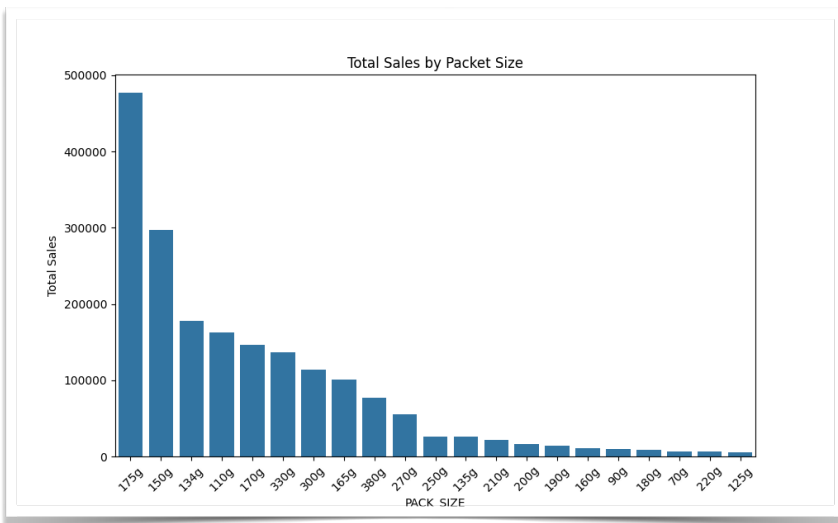**Explanation** : The bar chart illustrates total sales based on the packet size of products. It is evident that customers prefer larger pack sizes, with 175 grams being the most purchased. This indicates that people tend to buy products that last longer in a single purchase, favoring bulk sizes. Smaller pack sizes, on the other hand, have the lowest sales, suggesting they are less popular, possibly due to the convenience and value provided by larger packets

Explanation : This bar plot shows that the brand "*Kettle*" has the highest sales, indicating strong customer preference and possibly optimal placement or demand for this product. In contrast, "*Sunbites*" has the lowest sales, suggesting poor product placement, lower customer preference, or the need for better marketing strategies to improve its visibility and demand.
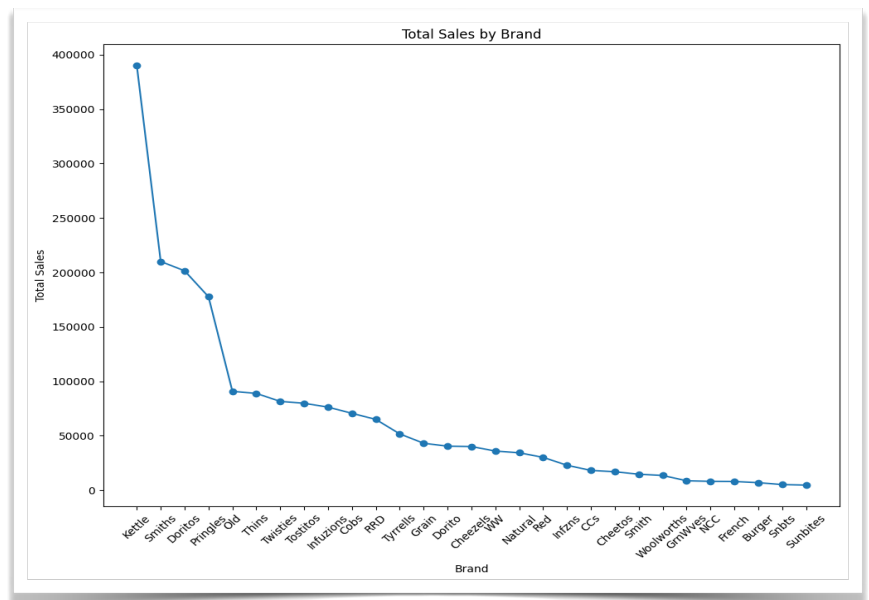

6. Sales Trends Over Time
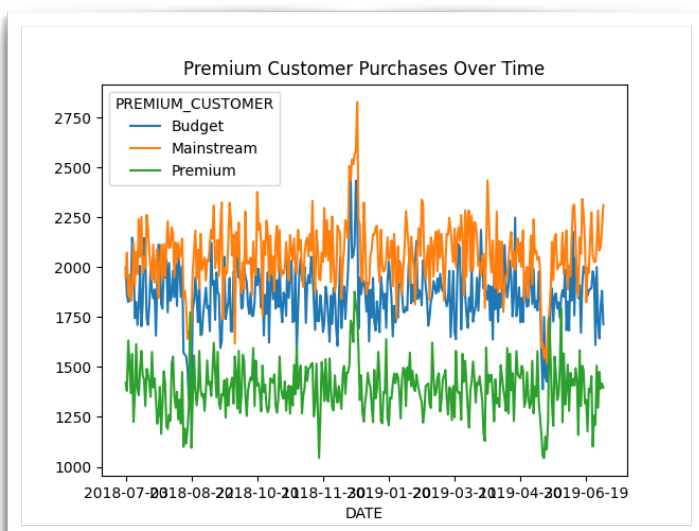
• **Explanation**: The graph shows mainstream customers drive high purchase activity, especially between November and January, indicating seasonal buying linked to holidays or promotions. Focusing on marketing to premium customers could further increase their purchase activity, while other segments remain balanced throughout the year.

7. Sales by Store.
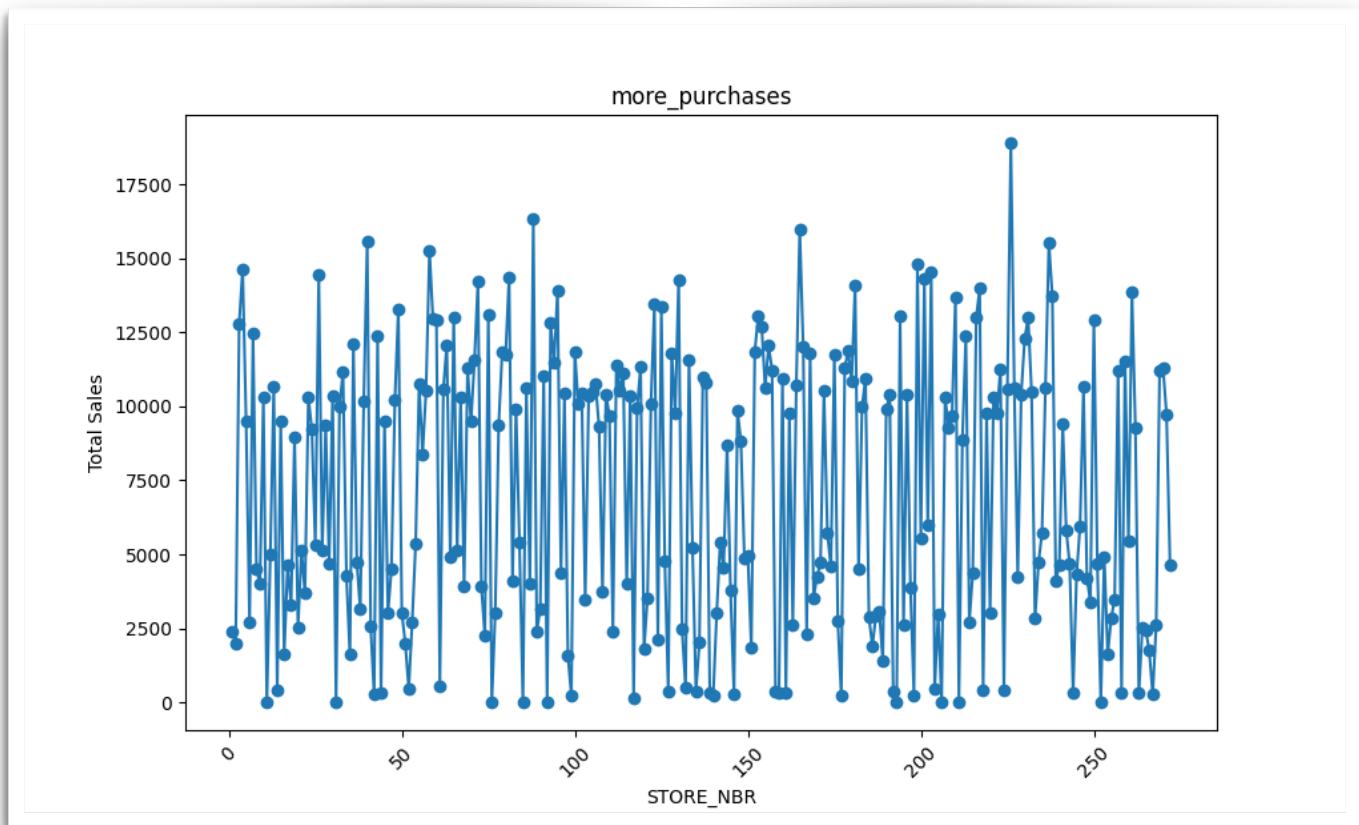
- **Explanation**: This analysis shows the performance of various stores in terms of sales. Stores ranging from 200 to 250 maintain consistent sales, with one store standing out with the highest sales, possibly due to its location in a high-traffic area or effective advertising. However, several stores in this range are underperforming, indicating a need for improvement. Strategies such as offering discounts, enhancing store marketing, or optimizing product placement could help boost sales in these lower-performing stores.

8. Suggestions for Strategy for Target Customer Segment:

o The data indicates that mainstream customers exhibit high levels of purchase activity, especially from November to January. Thus, it is crucial to concentrate on this market over the holidays.However, just 26% of overall purchases come from premium clients. This market has room to develop, therefore over the following six months, their sales contributions could rise with a focused marketing approach that draws in additional high-end clients.

o Product and Brand Prioritization: The best-selling brand is clearly "Kettle." Prioritizing this brand in marketing campaigns, promotions, and stock availability will optimize revenue.Conversely, "Sunbites" is not doing well, and there is room for improvement in terms of positioning or stocking strategy.

o Packet Size Strategy: According to the data, 175-gram packets are the most widely used size, with customers favoring lower sizes less. Customers tend to choose things that last longer per purchase, so it's advised to stock and promote larger pack sizes, especially 175-gram packets, to coincide with this purchasing tendency.

9. Conclusion

- High-demand products like Kettle should be positioned in strategic retail places, like end caps or next to the checkout area, to improve in-store advertising and product placement. This will stimulate impulsive purchases. In-store displays and eye-catching signage can help direct customers' attention to these top-selling goods. Furthermore, offering bundles that combine high-performing products, like Kettle, with complimentary goods, such drinks or dips, might entice buyers to make greater purchases, leading to higher basket sizes and higher overall sales.

- We can use targeted promotions to increase sales through individualized marketing for premium clients, who now account for only 26% of sales. Personalized emails or SMS offers that provide special discounts or incentives for making larger purchases can encourage this group to engage more and spend more money. Furthermore, creating a loyalty program for premium clients that awards points for every transaction or exclusive savings on their favored brands or items would promote recurring business and greater purchases, boosting their entire sales contribution.

- We should concentrate on the notable increase in sales that was seen between November and January in order to take advantage of seasonal and event-based marketing. During this time, early promotions with discounts, holiday-themed merchandise, and buy one, get one free deals can be introduced to encourage early shopping and increase sales. Additionally, targeted discounts and promotions should be made available in the middle of the year or throughout the summer to keep customers engaged during off-peak hours. This tactic keeps customers interested even outside of the busiest shopping seasons and helps maintain a balance in sales throughout the year.

- Organizing in-store activities, such as product tastings, freebies, or seasonal promotions, can help draw in local communities and increase foot traffic, which can assist improve sales in failing stores. These gatherings foster interaction and strengthen ties with clients. Furthermore, rearranging aisles and displays to optimize the store layout might increase high-demand products' visibility and accessibility to customers, resulting in greater sales. Better product placement together with community-focused events might help boost sales in these underperforming sites.

- Customer preferences should be the main focus when promoting larger pack sizes, since evidence indicates that larger packs such as 175 grams are more popular. The store ought to carry more of these sizes and advertise them more aggressively. Creating a subscription service for frequent customers or providing discounts for buying numerous large packs might help increase sales. Furthermore, offering bulk-buy discounts like "Buy 2 large packs, get 1 free" promotions can entice customers to make larger purchases in a single transaction, boosting overall sales and effectively moving more inventory.

- If you want to increase sales and get rid of excess stock, you should think about doing special promotions or discounts on underperforming products like "*Sunbites*". Better-selling brands may need to take precedence over "*Sunbites*" on the shelf if this tactic proves ineffective in decreasing their amount of space. To keep the product offers current and in line with consumer tastes, a product refresh or replacement strategy could also be used, bringing new or trending items. This could entail introducing new product lines, sizes, or flavors in an effort to pique consumer interest and boost overall sales results.

- Leveraging digital channels through targeted social media ads and email campaigns is crucial for improving digital marketing and online sales. Driving traffic both online and in-store can be accomplished by promoting limited-time online deals or special products that are only available to digital customers. Furthermore, consumers who would rather shop online but would like to pick up their products in-store will find it convenient because there is now a online ordering with in-store pickup option. This strategy increases overall shop visits and sales in addition to increasing digital interaction.