

Retail Analytics Report

To

Goals of the analysis:

1. Note any interesting trends and insights.
2. Deep dive into customer segments and recommend which segments should be targeted.
3. Determine if packet sizes are relative to sales.

Index:

1. *Introduction*
2. *Data Cleaning*
3. *Key Metrics*
4. *Analysis and Results*
5. *Conclusion*

Introduction:

A dataset of 264836 sales records was explored to analyze trends and patterns in customer purchasing behaviors to lead to insights that would act as a recommendation to focus on customer segment and specific products. Two files: transaction_data and purchase_behaviour were the data for this analysis.

Data Cleaning:

Both the files: transaction_data(file 1) and purchase_behaviour(file 2) were checked for any missing values and duplicated records. File 1 had 1 duplicated record which was removed from the dataset, whereas file 2 did not have any duplicated records. Both the files did not have any missing data.

Data that was not going to be used for calculations was converted into a string and category to maintain its uniqueness. Data in both the files that was converted to string: Date, Loyalty Card Number, Transaction ID, Product Number, Store Number.

Data from file 2 that was converted into category datatype includes: Lifestage and Premium Customer.

Since, one goal of the analysis was to check how pack size was related to sales, the pack size information was extracted from the Product Name column and was stored into two separate columns to separate the integer datatype from the string: Pack size weight and Pack size unit.

Once both the files were good to go, they were merged based on the common column: Customer Loyalty Card Number (LYLTY_CARD_NBR).

Key Metrics:

- A. General metrics were computed that would be used further in the analysis:
 - a. Total Sales: 1934408.99
 - b. Average sale per customer: 26.63
 - c. Average sale per transaction: 7.35
 - d. Average sale per store: 7111.79

Analysis

Sales Drivers:

The customers segment that brought in the most sales throughout the dataset are Older Families that fall under the budget category. Following them, Young Singles/Couples and Retirees categorized as mainstream drove the most sales. (see chart 1)

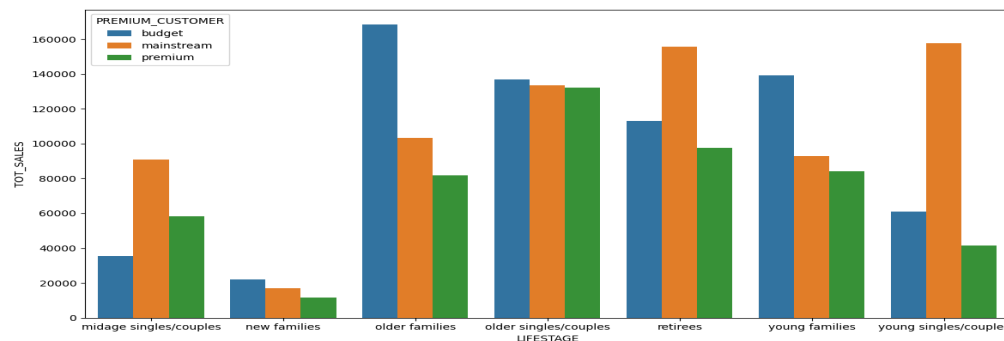


Chart 1

Whereas, on an average mainstream singles/couples brought in the most sales. (see chart 2)

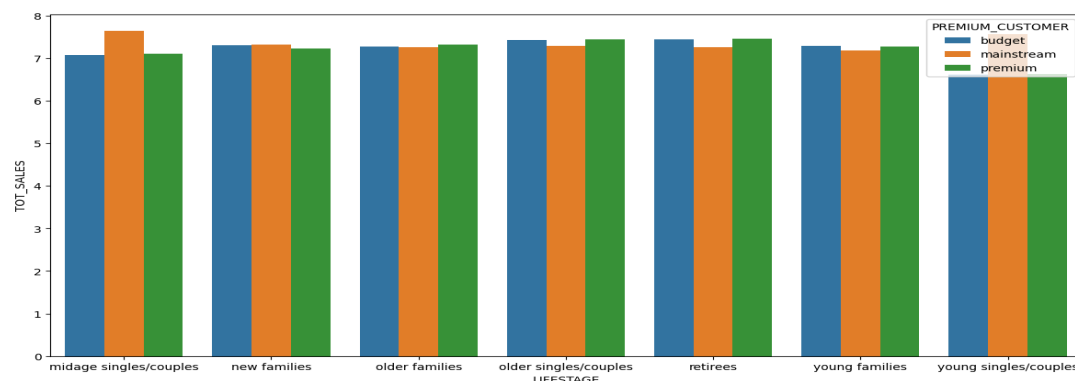


Chart 2

Best Selling Products:

Dorito Corn Chip 380g grossed the top selling product with a sale of 40,000\$. (see *Chart 3*)

Whereas, the least selling product was Woolworths Medium Salsa 300g with 4050.0\$ in sales. (see *Chart 4*)

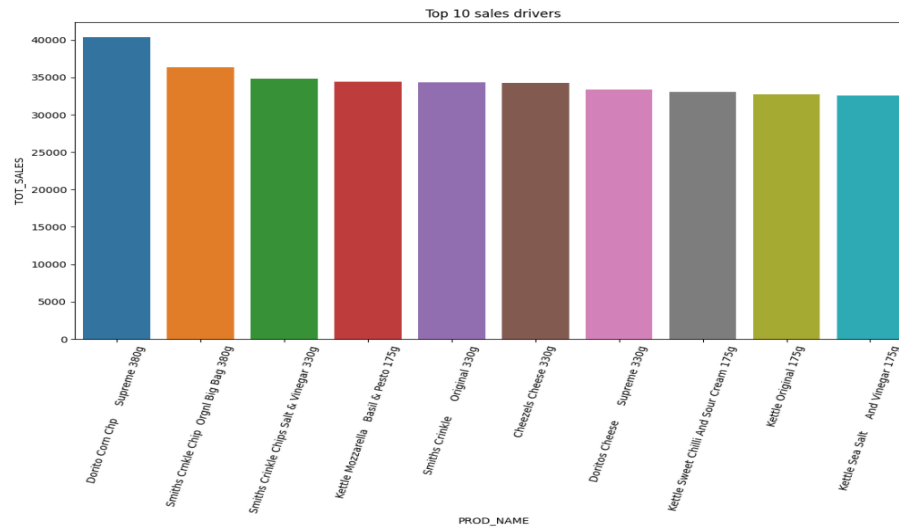


Chart 3

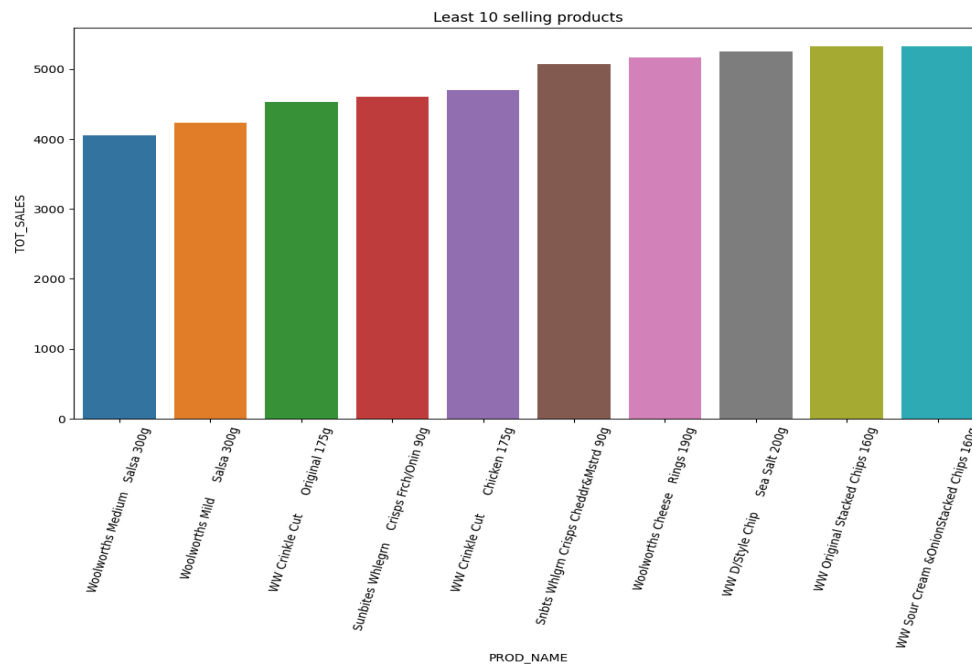


Chart 4

Package Size

Highest selling package size was found to be 175g that accounted for almost 500,000\$. (see Chart 5)

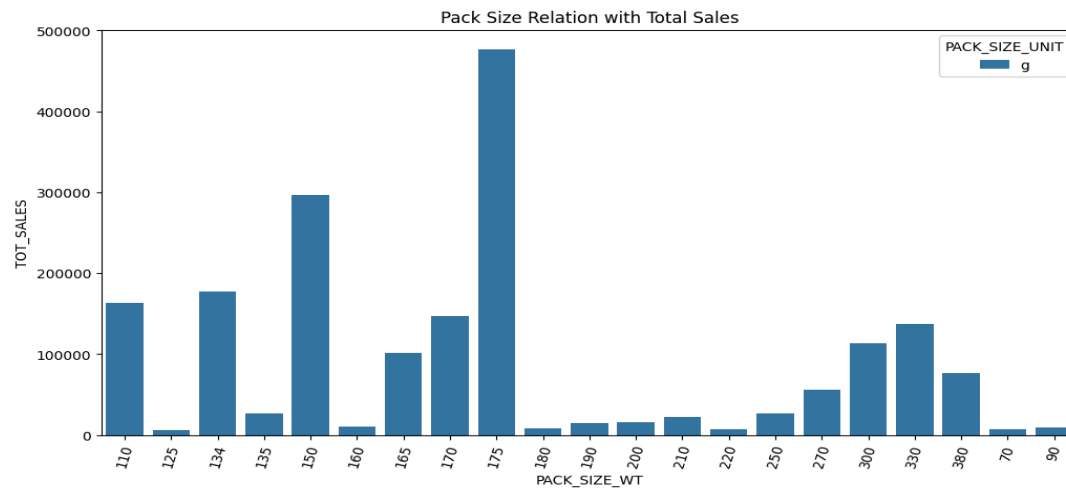


Chart 5

Conclusion:

Strategic Recommendations:

- A greater attention in terms of attraction can be given to Mid-age singles/couples.
- Since, Packages weighing 175g were significantly bought, Experiments can be performed where packages weighing 70g, 90g, 125g, 220g can be replaced with the same brand but either of 175g or 150g.

An interesting insight: Older families categorized as 'budget' accounted for the highest total sale but after calculating the average, this segment was just as good as others, which means older families visited the stores a lot more often than other segments.