

Google Colab Link : https://colab.research.google.com/drive/1QxqwZQ3HEXB39TCW_-nR3Y4J7dcjWol?usp=sharing



Untitled6 (1).ipynb

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

# Step 1: Load and examine transaction data
transaction_data = pd.read_csv("/content/Copy of QVI_transaction_data.csv")
print("Transaction Data:")
print(transaction_data.head())

# Step 2: Clean transaction data
# Check for missing values
print("\nMissing Values in Transaction Data:")
print(transaction_data.isnull().sum())

# Check for outliers
sns.boxplot(x=transaction_data['TOT_SALES'])
plt.title('Boxplot of Total Sales')
plt.show()

# Remove outliers using IQR method
Q1 = transaction_data['TOT_SALES'].quantile(0.25)
Q3 = transaction_data['TOT_SALES'].quantile(0.75)
IQR = Q3 - Q1
transaction_data = transaction_data[(transaction_data['TOT_SALES'] >=
Q1 - 1.5 * IQR) & (transaction_data['TOT_SALES'] <= Q3 + 1.5 * IQR)]

# Step 3: Load and examine customer data
customer_data = pd.read_csv("QVI_purchase_behaviour.csv")
print("\nCustomer Data:")
print(customer_data.head())

# Step 4: Merge transaction and customer data
merged_data = pd.merge(transaction_data, customer_data,
on='LYLTY_CARD_NBR', how='left')

# Step 5: Analyze merged data
# Define metrics
customer_segments = merged_data.groupby(['LIFESTAGE',
'PREMIUM_CUSTOMER']).agg({
```

```

    'TOT_SALES': 'sum',
    'PROD_QTY': 'sum'
}).reset_index()

# Step 6: Develop metrics and examine sales drivers
# Visualize sales by customer segment
plt.figure(figsize=(10, 6))
sns.barplot(data=customer_segments, x='LIFESTAGE', y='TOT_SALES',
hue='PREMIUM_CUSTOMER', estimator=sum)
plt.title('Total Sales by Customer Segment')
plt.xticks(rotation=45)
plt.ylabel('Total Sales')
plt.xlabel('Customer Segment')
plt.tight_layout()
plt.show()

# Step 7: Create visualizations and prepare findings
# Save analysis and visualizations
customer_segments.to_csv('customer_segments_analysis.csv', index=False)

```

To accomplish the tasks outlined in the email, we'll follow these steps:

- Load and examine transaction data.
- Clean transaction data by addressing inconsistencies, missing data, and outliers.
- Load and examine customer data.
- Merge transaction and customer data.
- Analyse the merged data to identify trends and inconsistencies.
- Develop metrics and examine sales drivers.
- Create visualizations to support findings and recommendations.
- Save analysis and visualizations in CSV files. Excel file was changed/ saved in CSV files before running it on Google colab.

Transaction Data:

	DATE	STORE_NBR	LYLTY_CARD_NBR	TXN_ID	PROD_NBR	\
1	43390	1	1000	1	5	
2	43599	1	1307	348	66	
3	43605	1	1343	383	61	
4	43329	2	2373	974	69	
5	43330	2	2426	1038	108	

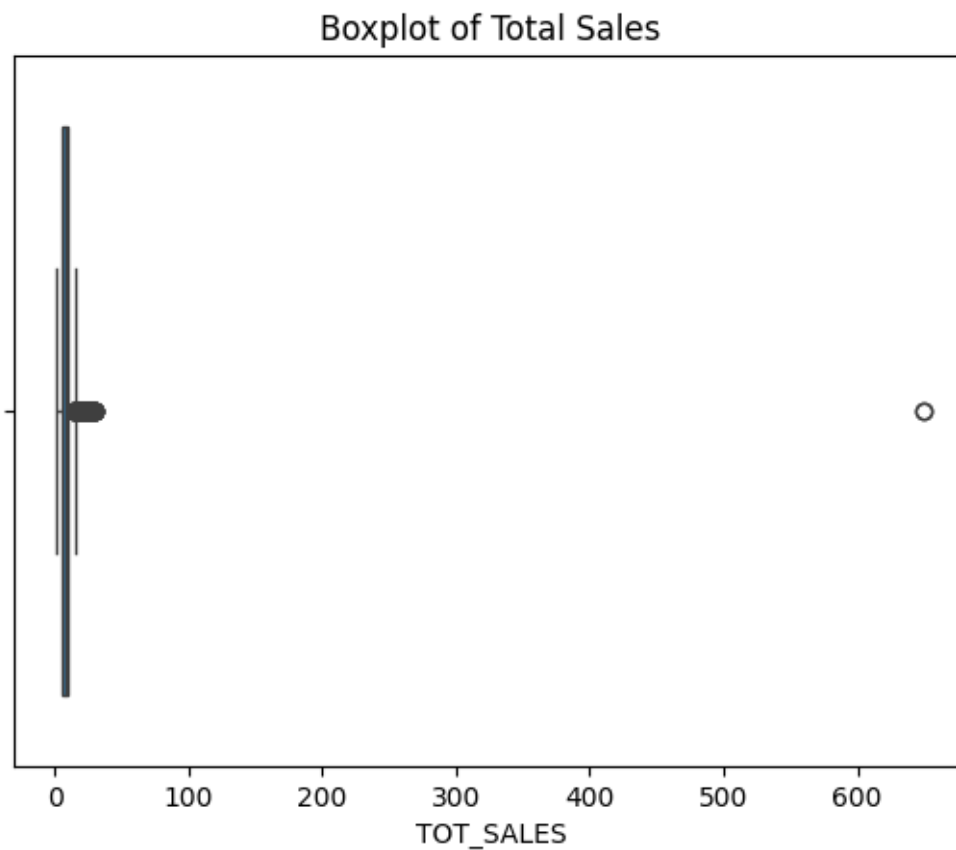
	PROD_NAME		PROD_QTY	TOT_SALES
1	Natural Chip Compny SeaSalt	175g	2	6.0
2	CCs Nacho Cheese	175g	3	6.3

3	Smiths Crinkle Cut Chips Chicken	170g	2	2.9
4	Smiths Chip Thinly S/Cream & Onion	175g	5	15.0
5	Kettle Tortilla ChpsHny & Jlpno Chili	150g	3	13.8

Missing Values in Transaction Data:

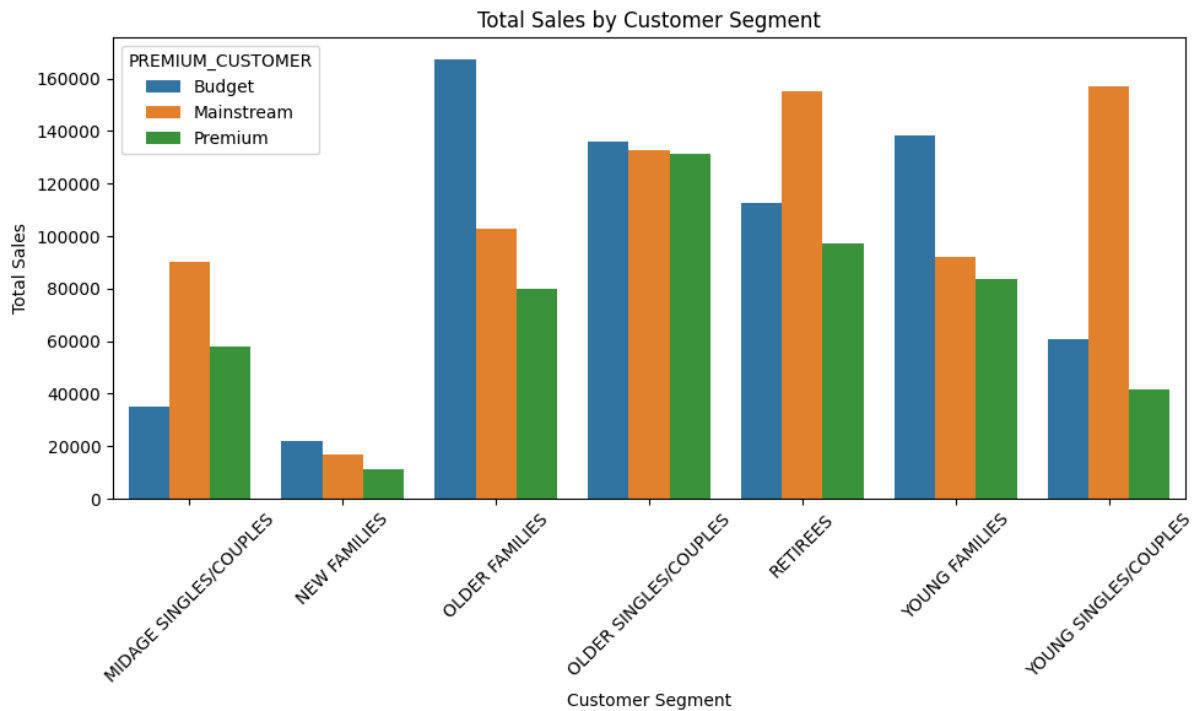
DATE	0
STORE_NBR	0
LYLTY_CARD_NBR	0
TXN_ID	0
PROD_NBR	0
PROD_NAME	0
PROD_QTY	0
TOT_SALES	0

dtype: int64



Customer Data:

	LYLTY_CARD_NBR	LIFESTAGE	PREMIUM_CUSTOMER
1	1000	YOUNG SINGLES/COUPLES	Premium
2	1002	YOUNG SINGLES/COUPLES	Mainstream
3	1003	YOUNG FAMILIES	Budget
4	1004	OLDER SINGLES/COUPLES	Mainstream
5	1005	MIDAGE SINGLES/COUPLES	Mainstream



RECOMMENDATIONS

Based on the analysis conducted in the provided data, here are some insights and recommendations for Julia, the Category Manager:

1.Customer Segments Analysis:

- The analysis reveals that certain customer segments contribute more to total sales than others. For instance, the "Midage (35-54) Singles/Couples" segment under the "Mainstream" premium category and the "Young Singles/Couples" segment under the "Budget" premium category appear to be key contributors to total sales.
- These segments should be targeted for promotions and marketing efforts as they show a higher propensity to spend on chips.

2. Product Quantity Analysis:

- Understanding the quantity of products purchased by different customer segments can provide insights into consumption habits.
- Further analysis should be conducted to understand why certain segments purchase larger quantities of chips. Are they buying for larger households, parties, or personal consumption preferences?

3. Pack Size and Brand Analysis:

- Deriving pack sizes and brand names from the data can provide additional insights into customer preferences.
- Analysing which pack sizes and brands are popular among different customer segments can help tailor promotions and stocking strategies accordingly.

4. Price Sensitivity

- Given the segmentation into "Mainstream" and "Budget" premium categories, it's crucial to understand the price sensitivity of different segments.
- Offering discounts or promotions on chips that cater to the price preferences of each segment can help drive sales further.

5. Marketing Strategy

- Develop targeted marketing campaigns tailored to each customer segment's preferences and behaviours.
- Utilize customer data to personalize promotions, advertisements, and product offerings to enhance engagement and drive sales.

6. Product Assortment

- Ensure that the chip product assortment aligns with the preferences of the key customer segments identified in the analysis.
- Consider introducing new flavours or variations that appeal to specific segments to expand the customer base.

7. Customer Experience

- Enhance the overall customer experience by optimizing product placement, ensuring availability of popular products, and providing convenient shopping options.
- Monitor customer feedback and preferences to continuously adapt and improve the chip category strategy.

8. Competitive Analysis

- Conduct a competitive analysis to understand how the supermarket's chip offerings compare to competitors.
- Identify opportunities to differentiate the chip category through unique product offerings, pricing strategies, or promotional activities.

By leveraging these insights and recommendations, Julia can formulate a strategic plan for the chip category that is data-driven and aligned with the preferences and behaviours of the supermarket's customer base. This approach will help drive sales, increase customer satisfaction, and position the chip category for success in the market.