Competitive analysis of Carboload category (PASTA) from brand (MUELLER)

James Li
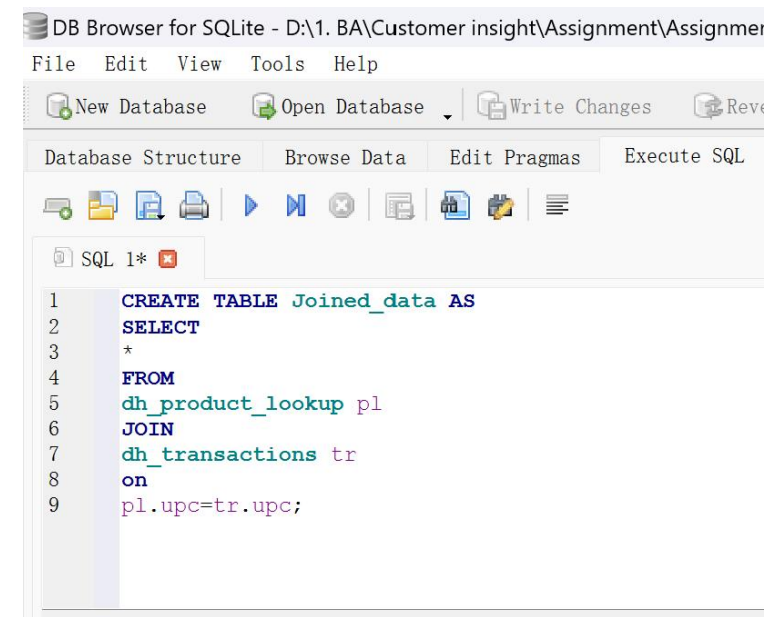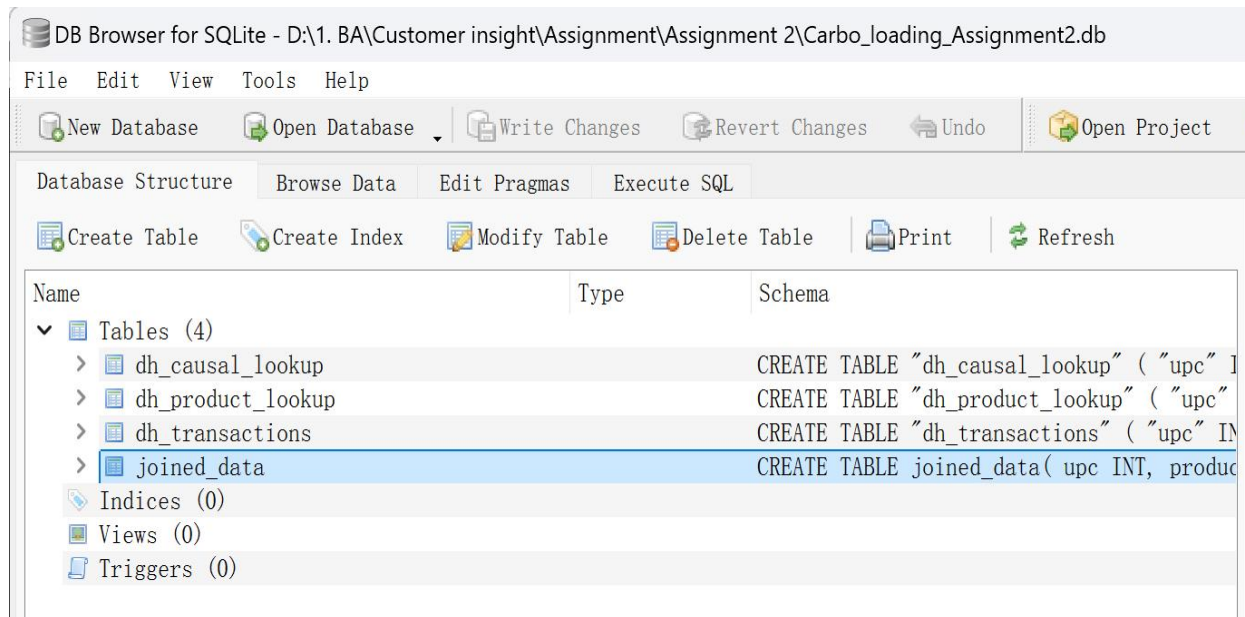
Content

1. Download and SQL

- Download the Dunnhumby carboload .csv files and combines into an SQL database.
- Write an SQL query to extract and save a .csv file containing relevant category data.


- exported file: dh_carboload raw_data.csv

2. Data Clean with Python and Jupyter

- Use Python and Jupyter notebooks to clean the data in file: dh_carboload raw_data.csv.

| 1	Duplicates	None
2	Missing	Products_size 43596
3	Outliers	dollar_sales drop where values are below zero
4	Columns to drop	time_of_transaction, geography, store, coupon, upc.1
5	Rows to drop	where dollar_sales<0
6	Feature extraction	where dolloar_sales<0
7	Cleaning	data_cleaned = data.copy() data_cleaned.fillna('MISSING') for col in data_cleaned.select_dtypes(include=['object']).columns: data_cleaned[col] = data_cleaned[col].str.strip() data_cleaned.drop(columns=['time_of_transaction', 'geography', 'store', 'coupon','upc:1'], inplace=True) data_cleaned = data_cleaned[data_cleaned['dollar_sales'] >= 0]
8	Filtering	data_cleaned = data.copy() data_cleaned.fillna('MISSING') for col in data_cleaned.select_dtypes(include=['object']).columns: data_cleaned[col] = data_cleaned[col].str.strip() data_cleaned.drop(columns=['time_of_transaction', 'geography', 'store', 'coupon','upc:1'], inplace=True) data_cleaned = data_cleaned[data_cleaned['dollar_sales'] >= 0]
9	Output	dh_pasta.to_csv('dh_pasta.csv', index=False) dh_pasta_sauce.to_csv('dh_pasta_sauce.csv', index=False) dh_syrups.to_csv('dh_syrups.csv', index=False) dh_pancake_mixes.to_csv('dh_pancake_mixes.csv', index=False)

3. Customer Dashboard -- Pasta

Quarter	Total_Num_Customers	Total_Num_Baskets	Total_Num_Stockcodes	Total_Quantity	Total_Revenue
1	137404	240167	349	353851	351531.55
2	141822	249890	347	350925	368235.29
3	138101	241644	349	350412	349341.68
4	128278	217567	337	316488	312454.34
5	141987	246161	333	356838	357253.45
6	149299	264182	325	388348	390429.49
7	147216	260395	322	397880	373448.65
8	139204	239957	315	363025	349849.26

Quarterly data of Pasta
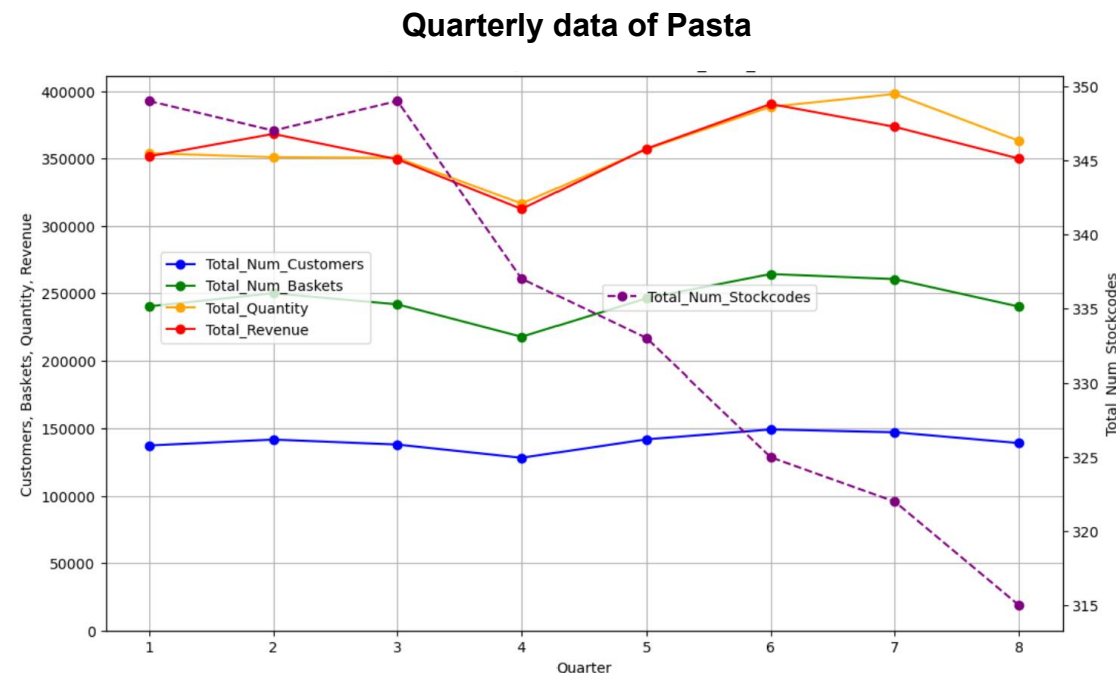
- From the curves in the chart, we can observe that "Total_Num_Customers," "Total_Num_Baskets," "Total_Quantity," and "Total_Revenue" are closely correlated, which is expected. There is a slight increase in "Total_Quantity" in Q7, while the other three variables decrease. This could be due to a promotion with reduced prices, where quantity sold increases but revenue declines.

- All four indicators show a decrease in Q4 but a recovery in Q5, likely due to seasonal factors, as we see a similar trend in Q8. Additionally, the decrease in "Total_Num_Stockcodes" indicates fewer new products were introduced.

3. Customer Dashboard -- Pasta

Quarterly data of Pasta based on per customer
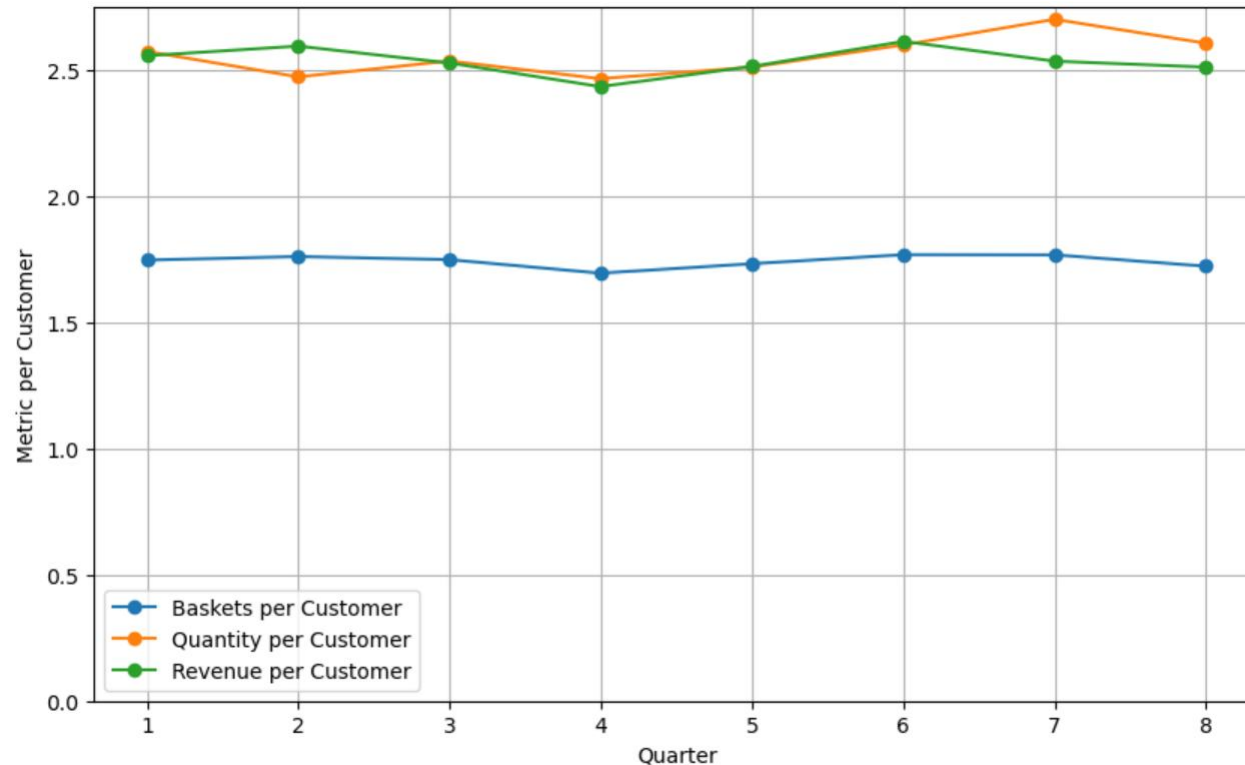

- The "Baskets per Customer" curve follows a similar trend to "Revenue per Customer," indicating that revenue per customer is related to how frequently customers make purchases.

- The "Quantity per Customer" curve shows a slightly different trend in Quarter 3 and 7. Customers purchase more from quantity side, while the revenue decreases, which indicates the price reduction or promotion.

3. Customer Dashboard -- Pasta

Quantity and Revenue per Basket
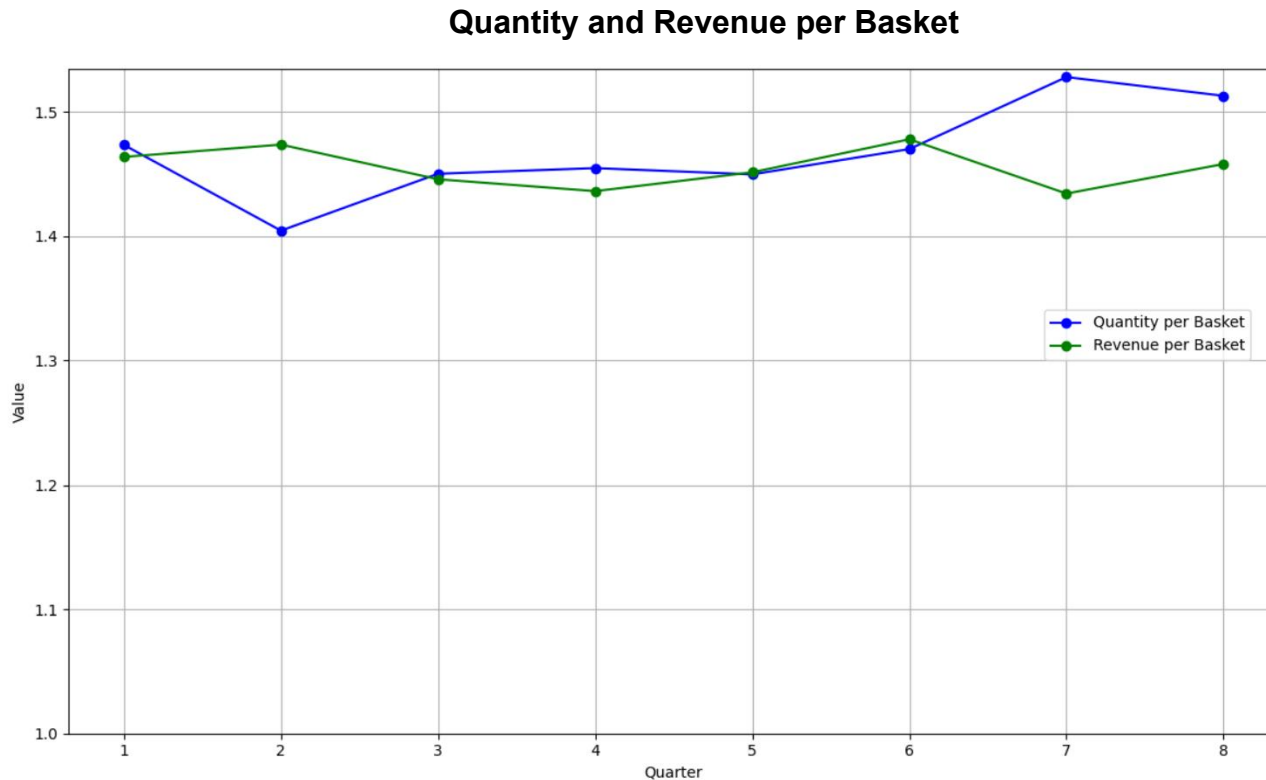
- The "Revenue per Basket" fluctuates slightly but generally remains around 1.45, indicating that customer spending per purchase has not changed significantly.

- On the other hand, "Quantity per Basket" shows relatively larger changes due to various factors. Seasonal price changes or promotions may lead to fluctuations in quantity, with customers purchasing more when prices drop, though this doesn't significantly impact the average revenue.

3. Customer dashboard -- Pasta

Percentage of different purchase class
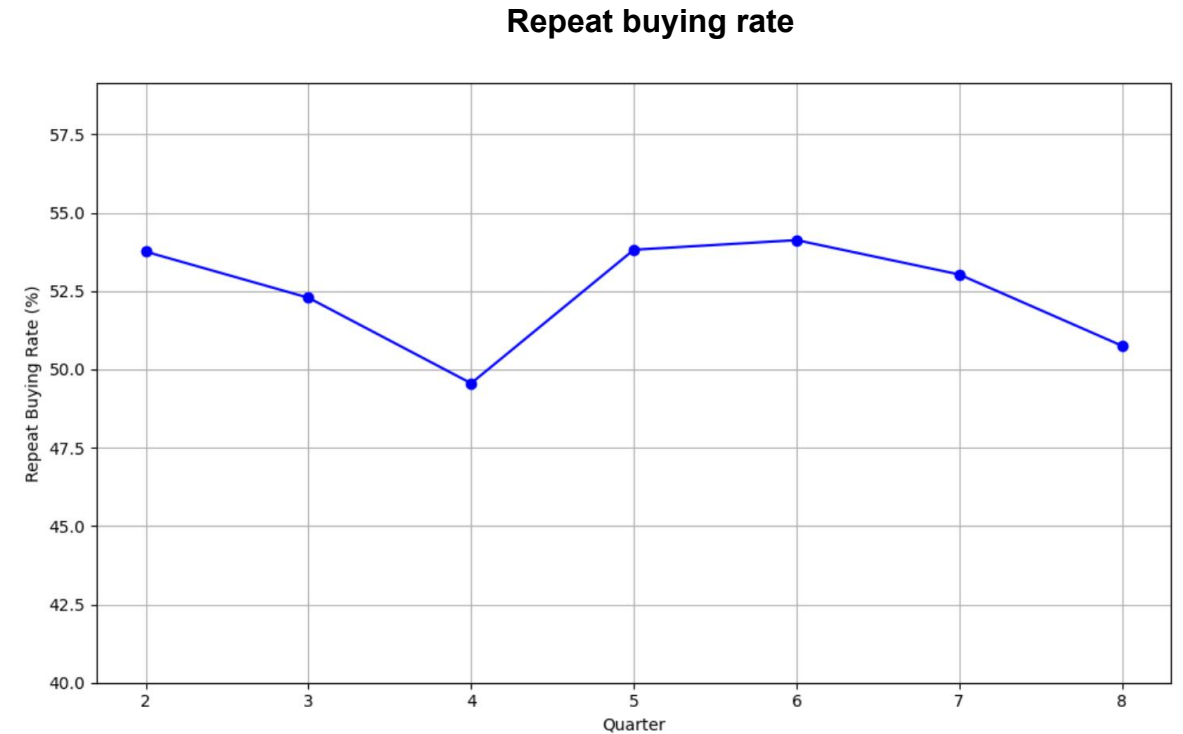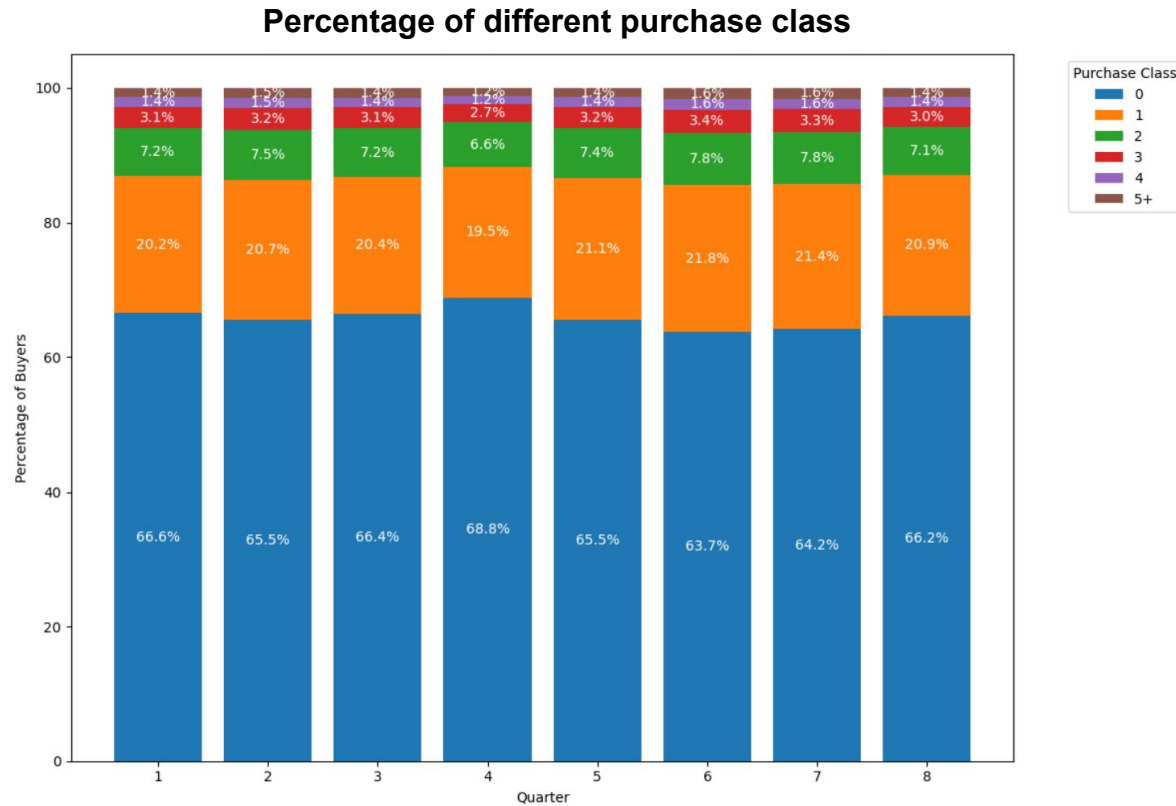
Repeat buying rate

- In Quarter 4, there is a slight increase in the "0 purchase class," while the "1 purchase class" and "2 purchase buyers" show a slight decrease. The repeat buying rate also declines in Quarter 4, which aligns with the trend observed in the purchase class chart. However, we see a revival in Quarter 5.

4. Customer dashboard -- Mueller

Mueller status in 8 quarters

Price of Mueller in 8 quarters
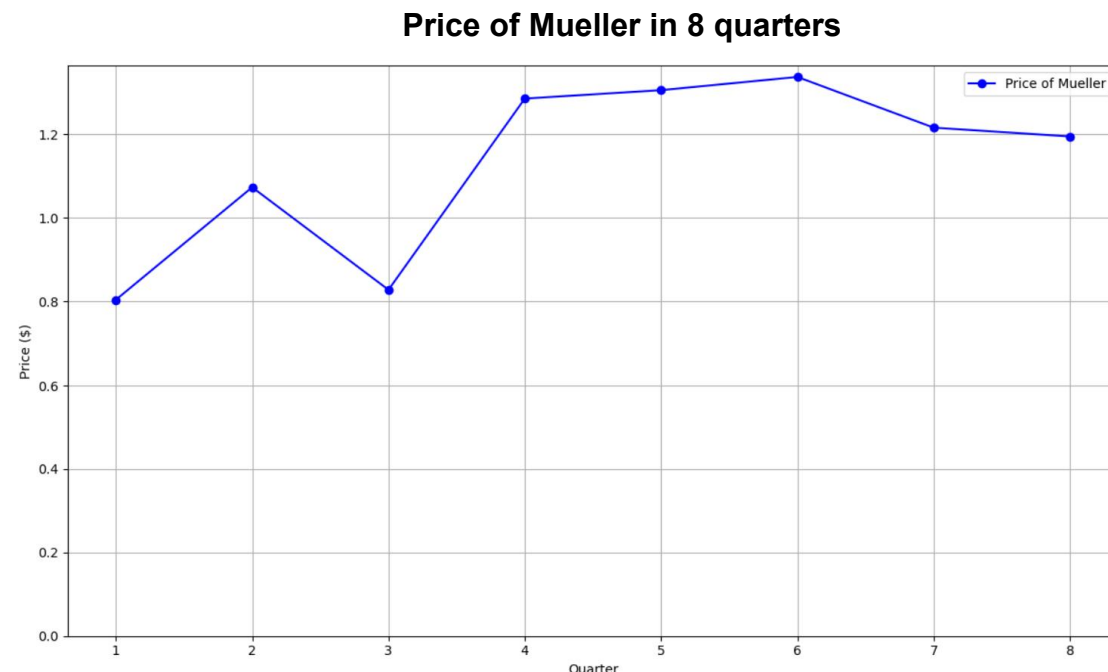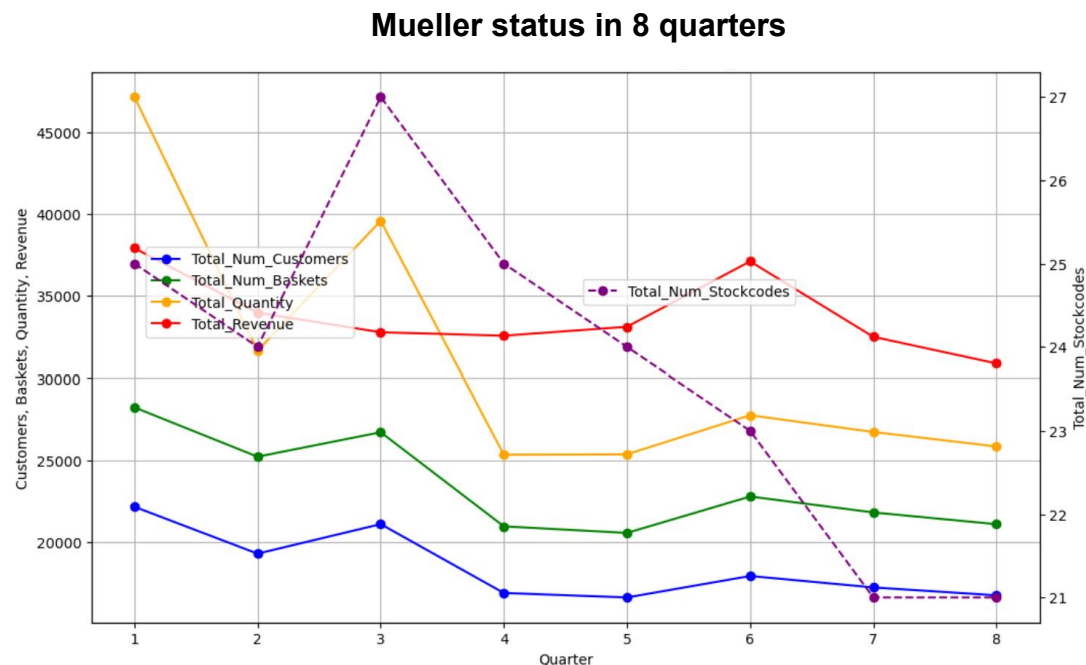
- Similar to the overall trend of category pasta, brand Mueller's indicators in the left chart show consistent patterns across different quarters. "Total_Quantity" for brand Mueller exhibits relatively larger fluctuations compared to "Total_Num_Customers," "Total_Num_Baskets," and "Total_Revenue."

- The trend in "Total_Quantity" for brand Mueller is closely linked to price changes, as shown in the right chart. A price decrease typically leads to an increase in quantity, and vice versa. We can observe that the significant price increase in Quarter 4 results in a drop in quantity, while revenue remains stable. Bigger size package could be the potential reason. However, the price decrease in Quarter 7 does not lead to an increase in quantity.

4. Customer dashboard -- Mueller

Category Average Price vs Price of Mueller
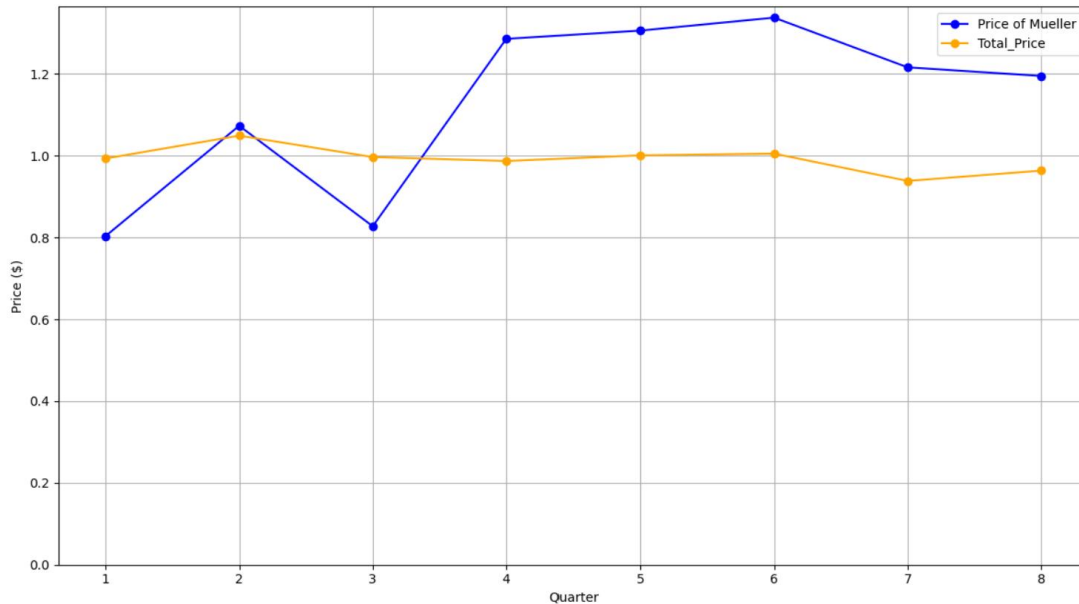
Mueller status in 8 quarters
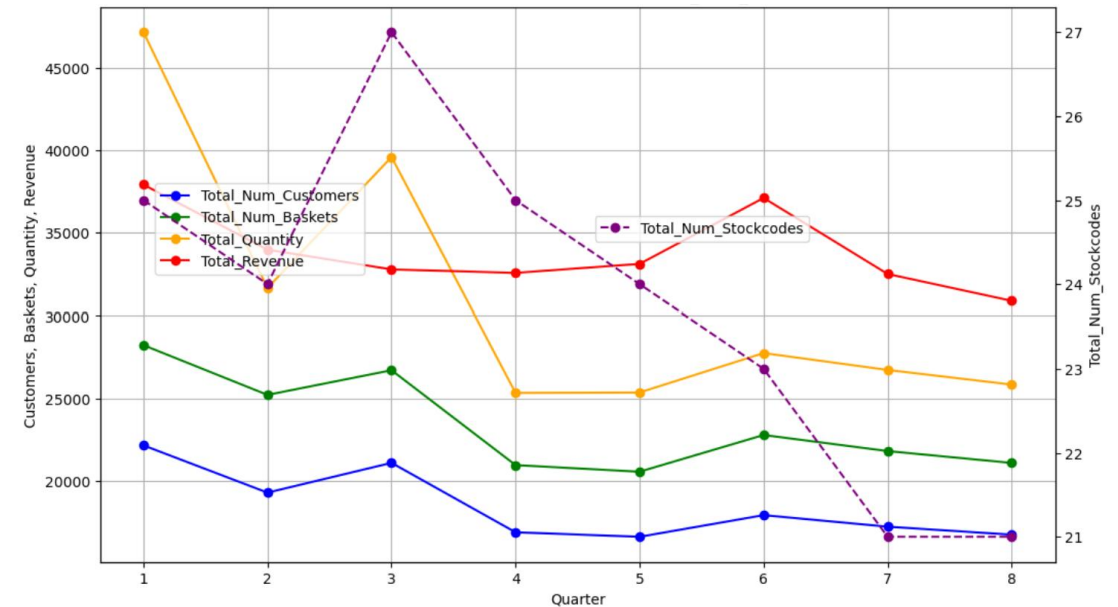
- From the left chart, we observe that Mueller's price trend generally aligns with the overall category average price, with the exception of Quarter 4. During this period, Mueller significantly increased its prices while the category average experienced a slight decline. Remarkably in the right chart, Mueller's revenue results for Quarter 4 were positive, showing only a minor decrease in revenue despite the substantial price increase, even though the quantity sold decreased.

- This stable momentum in sales revenue continued from Quarter 4 to Quarter 8, indicating that Mueller has successfully implemented its new promotional policy (potentially featuring higher prices alongside larger package sizes) while maintaining a loyal customer base without any revenue decline.

4. Customer dashboard -- Mueller

Purchase class of Mueller in different quarters
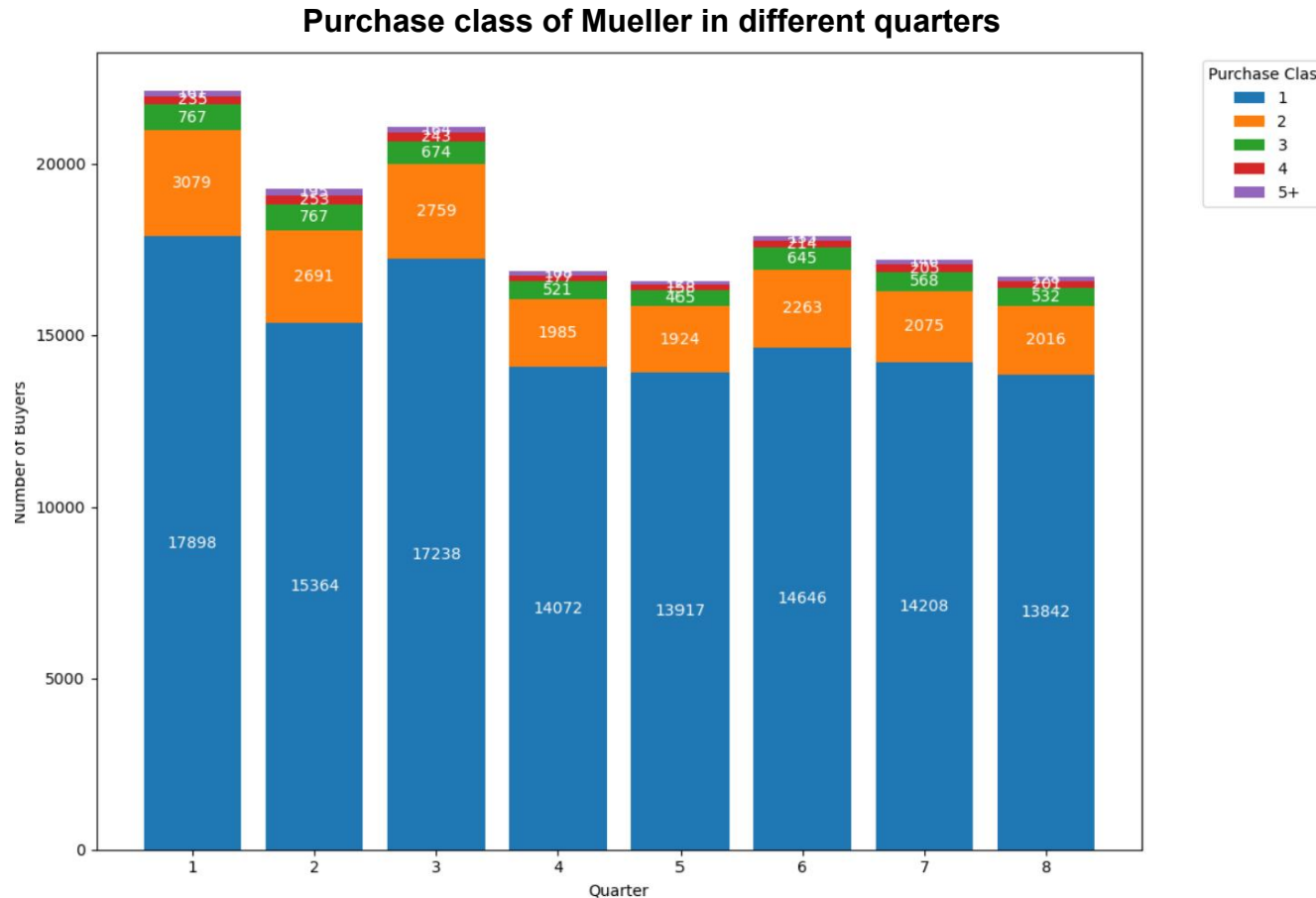
- By breaking down the "Total_Num_Customers" in previous page, we can observe the distribution of different purchase classes across various quarters in the left chart.

- From the left chart, we can see that the highest number of customers occurs in Quarter 1, which, as discussed earlier, also results in the highest revenue

- Big decrease of all purchase classes in Quarter 4 caused by new promotion strategy of Mueller.

4. Customer dashboard -- Mueller

Position of Mueller in all the brands in percentage

Legend:
- Percentage of Total_Num_Customers_mueller
- Percentage of Total_Num_Baskets_mueller
- Percentage of Total_Quantity_mueller
- Percentage of Total_Revenue_mueller
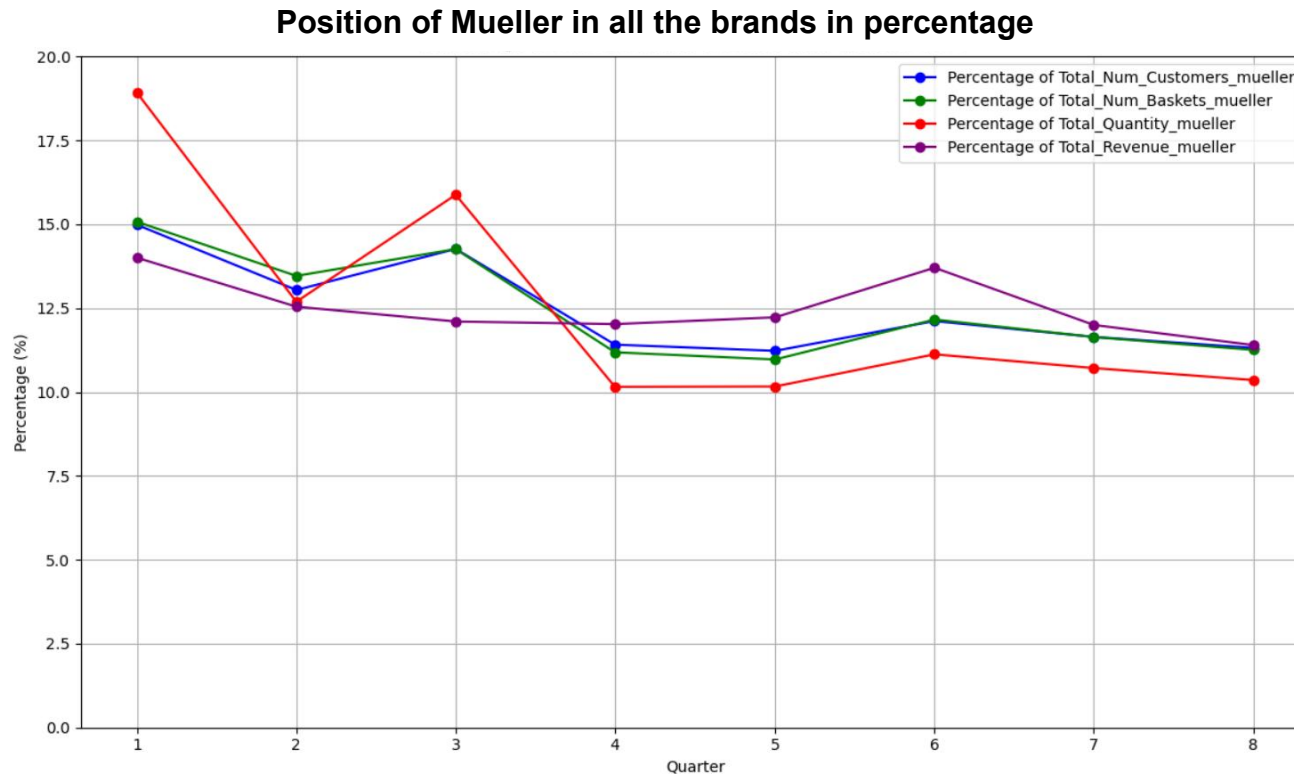
- The percentage of total quantity for brand Mueller experiences a significant decrease in Quarter 4, driven by the price increase. This also leads to a decrease in the percentage of total customers and total baskets across all brands.

- However, when examining the percentage of total revenue for Mueller, we find that there is almost no change in Quarter 4. This suggests that the price increase did not result in a revenue decline for Mueller among all brands. If we consider the potential impact on profit in Quarter 4, it's likely that the stable revenue alongside reduced quantity could indicate higher profitability.

5. Customer dashboard -- Quarter 8

Quarter 8

Legend: Total_Num_Customers, Total_Num_Baskets, Total_Quantity, Total_Revenue
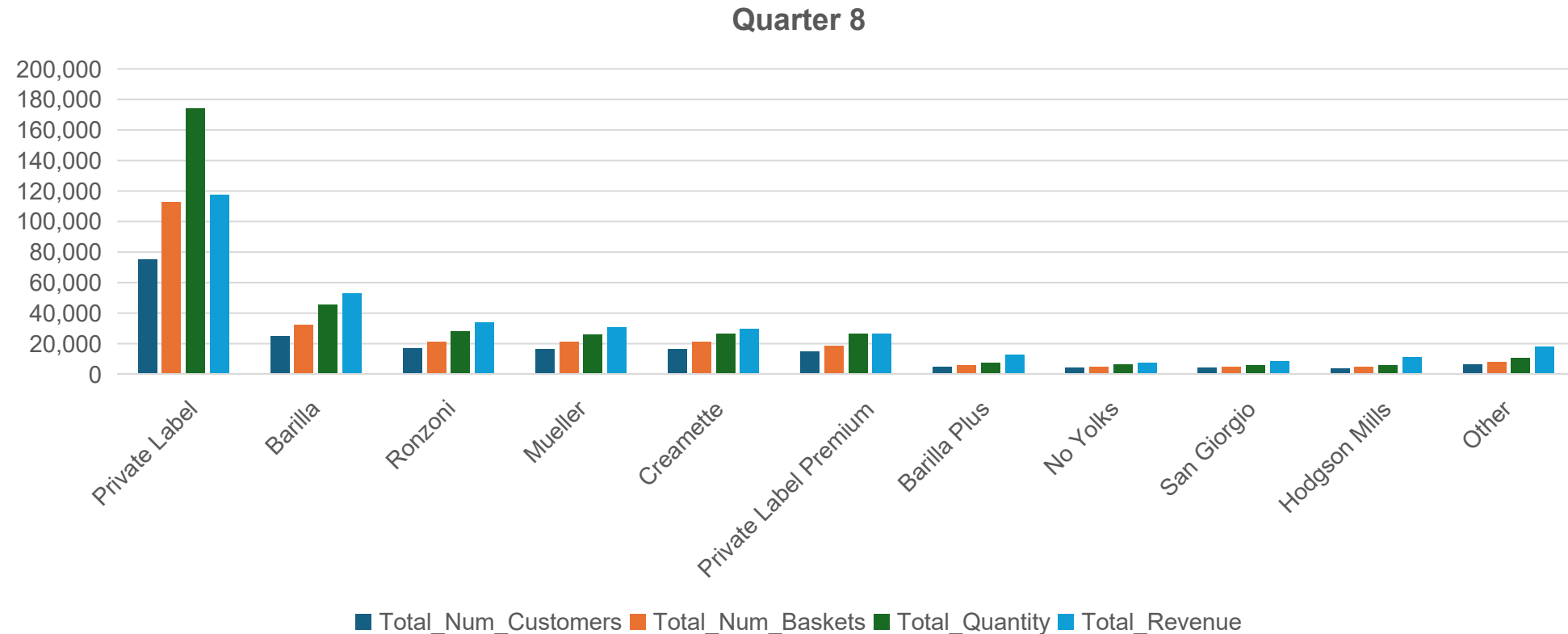
- For the four indicators, nearly all brands follow a similar trend, except for "Private Label."

- "Private Label" shows relatively lower total revenue, which could be due to its lower pricing. Frequent price promotions may also be a common strategy for this brand.

- "Mueller" is ranking number 4 from overall situation.

5. Customer dashboard -- Quarter 8

b and w of different brands in quarter 8

b and w correlation

$R^2 = 0.861$
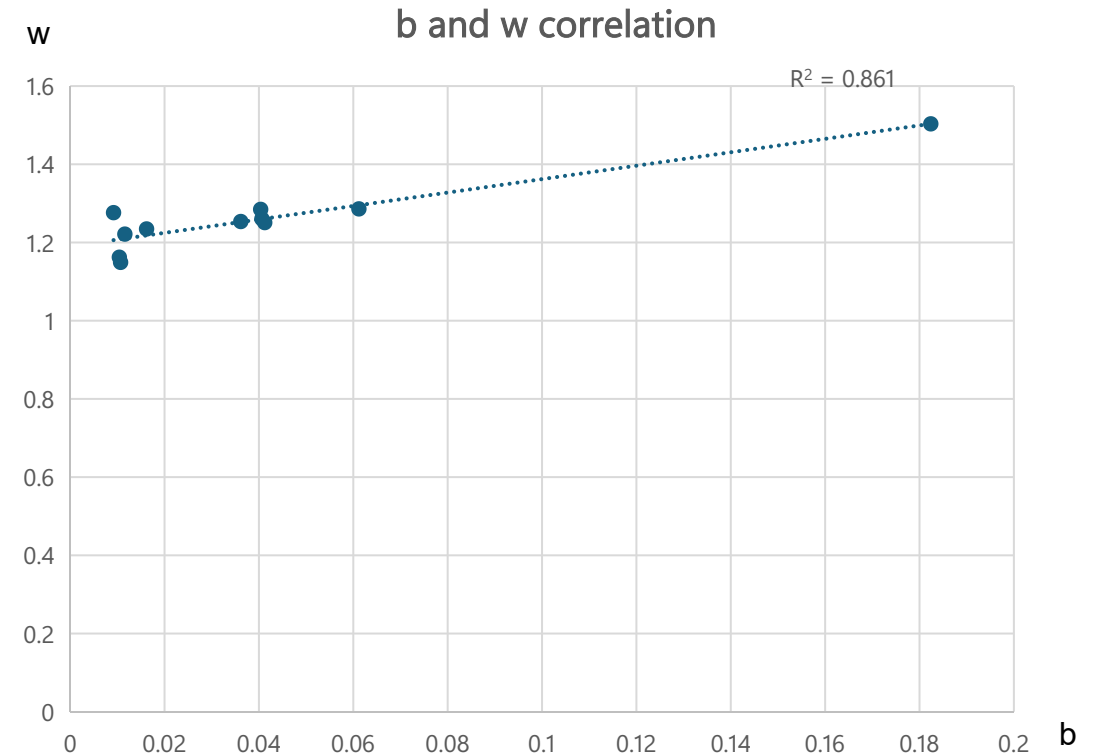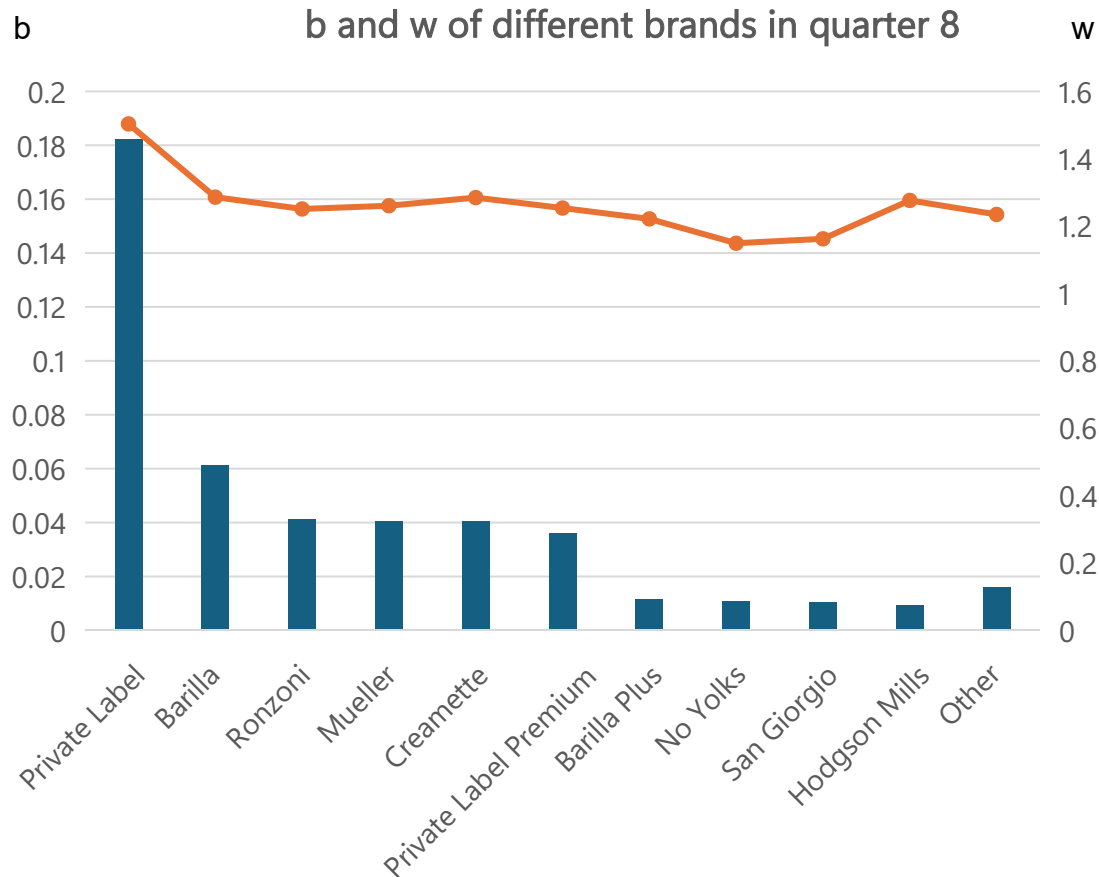
- From the output file for Quarter 8 generated in Python, we calculated the "b" and "w" values for each brand. "Private Label" has the highest penetration rate at 18% and the highest buying rate at 1.5. Muelle is ranking number 4 with b of 4% and w of 1.3.

- The "b" value is strongly correlated with the "w" value, exhibiting a high R-squared value. A high penetration rate or market share for a brand results in a greater influence in the market, which in turn leads to a higher buyer rate.

6. NBD Dirichlet -- Quarter 8

DATA ENTRY

Product Field and Brand Information

Product	B	W	P0	M	T=
Product	34%	1.7	66%	0.6	1

Input Data

Brand	b	w	Mkt Share Use to Est
1 Private Label	18%	1.5	Y
2 Barilla	6%	1.3	Y
3 Ronzoni	4%	1.3	Y
4 Mueller	4%	1.3	Y
5 Creamette	4%	1.3	Y
6 Private Label Premi	4%	1.3	Y
7 Other	2%	1.2	Y
8 Barilla Plus	1%	1.2	Y
9 No Yolks	1%	1.1	Y
10 San Giorgio	1%	1.2	Y
11 Hodgson Mills	1%	1.3	Y

- Here we input the calculated b and w, and also B and W.

6. NBD Dirichlet -- Quarter 8

Output of NBD Dirichlet

Model Fitting Statistics

K	.628	A	0.9	s^	1.48

Penetration			Purchases Per Buyer		
AVE%	1.6%	✓	AVE%	4.0%	✓
.r	1.00	✓	.r	0.93	✓
MAD%	3.0%	✓	MAD%	4.4%	✓
MAPE	4.3%	✓	MAPE	4.6%	✓

Number of Tests Passed	8	out of	8

- The model fits well, both in terms of predicting household penetration and the number of purchases per buyer.

- The low error percentages (AVE%, MAD%, MAPE) and high correlation values (.r) indicate that the model provides accurate predictions.

- Additionally, the fact that it has passed all tests signifies robustness in the model's overall performance.

6. NBD Dirichlet -- Quarter 8

Output of NBD Dirichlet

No.	Brands	Penetration		% of Buying		Purchases Per Buyer			Share of category requirements	100% Loyal		Repeat Buying
				r=1	r=5+	of the Brand		of the Category		%	Rate	
		O	T	T	T	O	T	T	T	T	T	T
1	Private Label	18%	19%	70%	1%	1.5	1.5	1.9	77%	71%	1.4	45%
2	Barilla	6%	6%	78%	1%	1.3	1.3	2.0	65%	58%	1.3	35%
3	Ronzoni	4%	4%	79%	1%	1.3	1.3	2.1	63%	56%	1.3	33%
4	Mueller	4%	4%	79%	1%	1.3	1.3	2.1	63%	56%	1.3	33%
5	Creamette	4%	4%	79%	1%	1.3	1.3	2.1	63%	56%	1.3	33%
6	Private Label Premium	4%	3%	79%	1%	1.3	1.3	2.1	63%	56%	1.3	33%
7	Other	2%	2%	80%	1%	1.2	1.3	2.1	61%	54%	1.3	31%
8	Barilla Plus	1%	1%	80%	0%	1.2	1.3	2.1	61%	54%	1.3	31%
9	No Yolks	1%	1%	80%	0%	1.1	1.3	2.1	61%	54%	1.3	31%
10	San Giorgio	1%	1%	80%	0%	1.2	1.3	2.1	61%	54%	1.3	31%
11	Hodgson Mills	1%	1%	80%	0%	1.3	1.3	2.1	61%	54%	1.3	31%

- For b and w, there is no big difference between the observed and theoretical results.
- For % of buying, all brands have a high percentage >=70% for 1 purchase class and small percentage <=1% for 5+ purchase class. So the purchase class overall is not high for the product pasta. And the purchases per buyers either for the brand or for the category is also around 1-2, which is not big.
- The "Private Label" brand dominates in both penetration and buyer engagement, indicating strong consumer loyalty and preference. Other brands exhibit a significant gap compared to "Private Label," with minimal differences among themselves across all indicators.
- Mueller is ranking number 4 overall, and its potential competitors are Barilla, Ronzoni, Creametter and Private Label Premium.

7. Summary

Insights:

- **Private Label** brand dominates both in penetration and buyer engagement, indicating strong consumer loyalty and preference.

- **Mueller** and similar brands have competitive buying percentages, suggesting potential for growth with effective marketing strategies.

- **Lesser-known** brands struggle with both penetration and purchase frequency, indicating opportunities for strategic improvements.

- All brands experience seasonal changes and dynamic competition within the brands of the category.

- The "b" value is strongly correlated with the "w" value, exhibiting a high R-squared value.

- For brand Mueller, in quarter 4 it successfully implemented its new promotional policy (potentially featuring higher prices alongside larger package sizes) while maintaining a loyal customer base without any revenue decline.

- From NBD Dirichlet, Mueller's potential competitors are Barilla, Ronzoni, Creametter and Private Label Premium.

7. Summary

Recommendations:

- Focus marketing efforts on increasing brand visibility and engagement for lower-performing brands.

- Explore customer loyalty programs or promotional strategies for brands like Mueller to enhance purchase frequency and penetration.

- Implementing the right pricing policy during seasonal changes and price competition is essential for defending market position and maintaining profitability.

- While Mueller's new promotion strategy in Quarter 4 has proven successful, it doesn't imply that every brand can achieve similar results with the same approach. Brand loyalty and targeted promotions during this period are crucial factors in this success, and it's worth noting that Mueller also lost some customers during this time. Strategies to regain these customers should also be considered.

- Mueller should closely monitor the pricing and promotional activities of potential competitors, such as Barilla, Ronzoni, Creamette, Private Label Premium, and the dominant Private Label brand in the market. NBD Dirichlet is a good tool for supporting the competitor analysis. Meanwhile, quick and proactive responses to the competition, along with strategies to differentiate itself from these brands, will be key in attracting new customers and enhancing repeat purchases and customer loyalty.

Thank you