

Shine Case - Marketing Analysis

Sarah, Mengxi, Yiyang, Ninh





TABLE OF CONTENTS

01

Introduction

Analysis

02

03

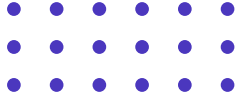
Recommendation

Conclusion

04



01



Introduction



Business Problem

Shine Cosmo-Cosmetics Inc., a medium-sized cosmetics company, is considering incorporating business analytics into their **2018 marketing strategy** to address slowed sales growth amidst online market challenges.

Timeline approaching for marketing plan.

With the use of analytics, the company will be able to make informed decisions for years to come as it grows in revenue and customer satisfaction.

Business Problem



Approach 1

1: Targeting Customers: Shine can focus on a young demographic cohort and offer personalized coupons to incentivize repeat purchases.

2: Alternatively, they can target new customers and offer deeper discounts to attract them to the brand.

Approach 2

1: Focusing on Merchandise: Shine can select specific items or bundles and heavily promote them to make them more appealing to customers.



02



Analysis

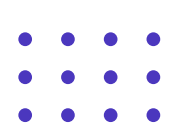


Data Exploration and Preprocessing

Data Loading: Two datasets, 'Shine.Case.Data_2022S.xlsx' and 'SKU_list.xlsx', are loaded into pandas dataframes.

	SKU	Regular retail price	Cost
Facial Care Series	FCS48020	79.99	48
	FCS48021	56.99	34
	FCS48022	94.99	60
	FCS48023	99.99	65
	FCS48024	59.99	34
	FCS48025	74.99	45
	FCS48026	56.99	35
	FCS48027	94.99	57
	FCS48028	69.99	42
	FCS48029	44.99	20
	FCS48030	74.99	45
	FCS48031	38.99	25
	FCS48032	59.99	36
	FCS48033	94.99	59
Skin Care Series	SCS72041	69.99	40
	SCS72042	44.99	32
	SCS72043	59.99	35
	SCS72044	59.99	35
	SCS72045	99.99	63
	SCS72046	99.99	60
	SCS72047	74.99	45
	SCS72048	69.99	42
	SCS72049	79.99	48

sku	Retail price	Cost
FCS48020	79.99	48
FCS48021	56.99	34
FCS48022	94.99	60
FCS48023	99.99	65
FCS48024	59.99	34
FCS48025	74.99	45
FCS48026	56.99	35
FCS48027	94.99	57
FCS48028	69.99	42
FCS48029	44.99	20
FCS48030	74.99	45
FCS48031	38.99	25
FCS48032	59.99	36
FCS48033	94.99	59
SCS72041	69.99	40
SCS72042	44.99	32
SCS72043	59.99	35
SCS72044	59.99	35
SCS72045	99.99	63
SCS72046	99.99	60
SCS72047	74.99	45
SCS72048	69.99	42



Data Exploration and Preprocessing

1. **Transaction ID Assignment:** Unique transaction IDs are assigned to each record in the dataset.
2. **Data Merging:** The datasets are merged based on the SKU column to provide comprehensive information for each transaction.
3. **Discount Conversion:** The discount column, initially in string format with a trailing '%', is converted into numeric values for further calculations.
4. **Actual Price Calculation:** An 'actual_price' column is added, which represents the product price after applying the discount. The values are rounded to two decimal places for accuracy.
5. **Profit Calculation:** A 'profit' column is created, representing the difference between the actual price and the cost for each SKU. The values are rounded to two decimal places.
6. **Total Profit Calculation:** A 'total_profit' column is created to account for transactions with quantities greater than 1. It multiplies the quantity by the profit for each SKU.
7. **Spend Calculation:** A 'spend' column is added, which multiplies the quantity by the original price for each SKU, providing the total expenditure for each transaction.

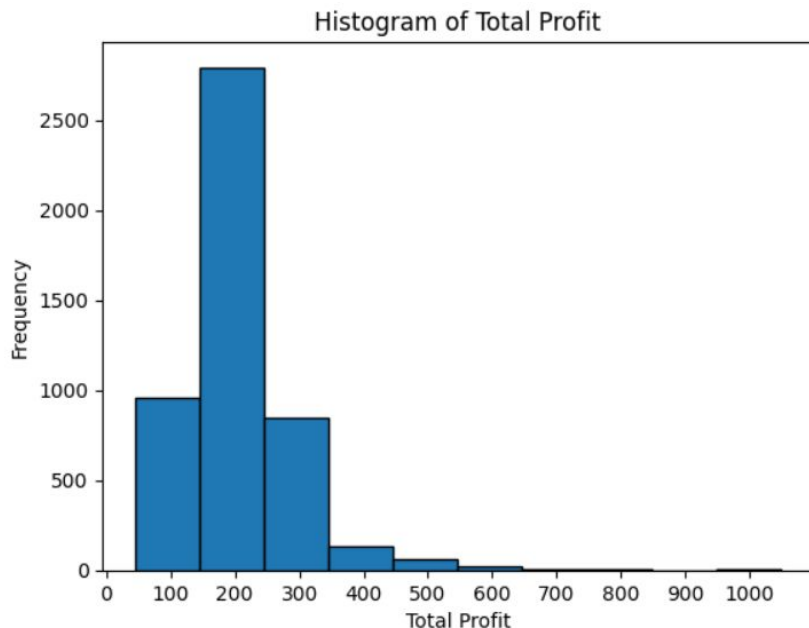


Get Customer data

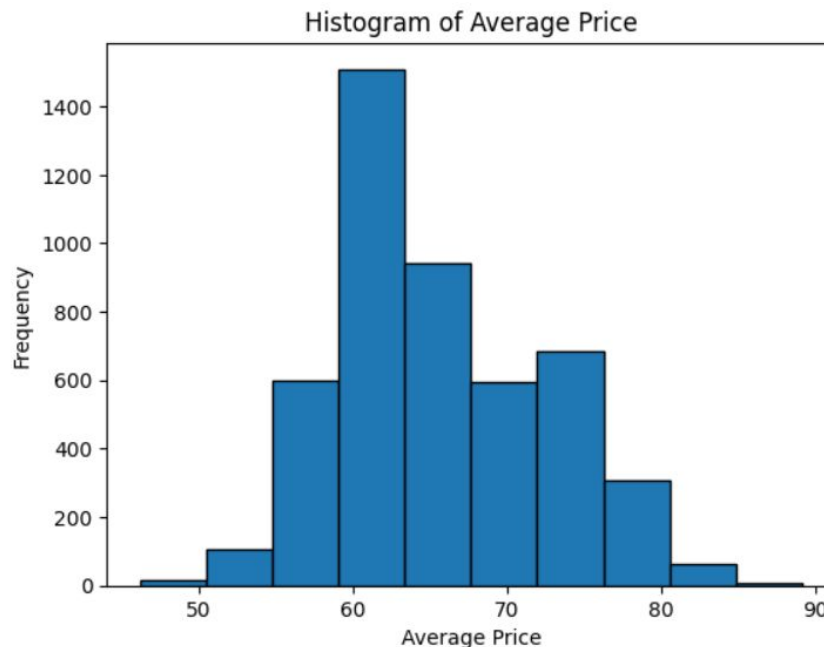
	customer_id	id	age	gender	tenure	average_price	average_discount	total_quantity	total_return	total_profit	total_spend	last_order_date	number_orc
0	1	110010	22	F	2	61.906667	0.050000	12	0	273.98	742.88	2017-06-19	
1	2	110039	34	F	10	70.740000	0.062500	8	0	198.52	565.92	2017-06-03	
2	3	110114	27	F	0	60.212222	0.118750	9	1	159.91	541.91	2017-05-30	
3	4	110125	26	F	0	55.656667	0.144444	9	0	129.16	500.91	2017-05-22	
4	5	110127	23	F	0	59.890000	0.045000	10	0	216.45	598.90	2017-06-23	
5	6	110152	27	F	1	64.808182	0.172727	11	1	145.64	712.89	2017-05-11	
6	7	110153	28	F	0	62.990000	0.014286	7	0	177.23	440.93	2017-04-15	
7	8	110220	32	F	7	72.156667	0.062500	12	1	268.88	865.88	2017-07-19	
8	9	110232	25	F	0	56.740000	0.043750	8	0	168.92	453.92	2017-06-12	
9	10	110256	29	F	2	65.790000	0.110000	5	0	102.10	328.95	2017-06-05	



Visualize Customer Data

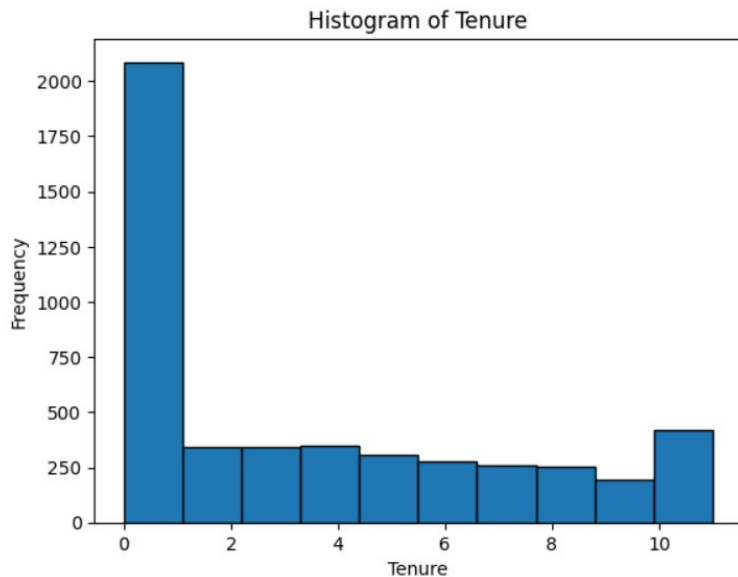


The typical profit the company made from each customer is around \$200, with the majority of profits falling within the range of \$100 to \$300.

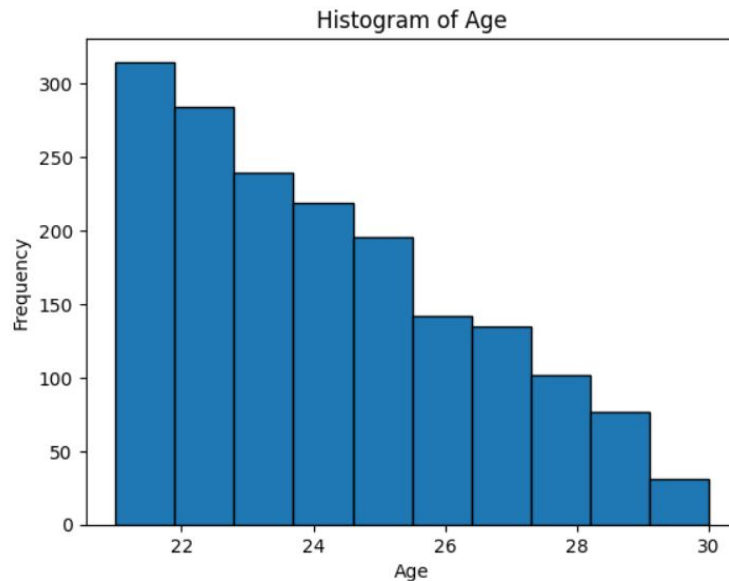


The average price (spending) of each customer are around \$55 to \$80, where \$60 to \$65 were spent most.

Visualize customer data

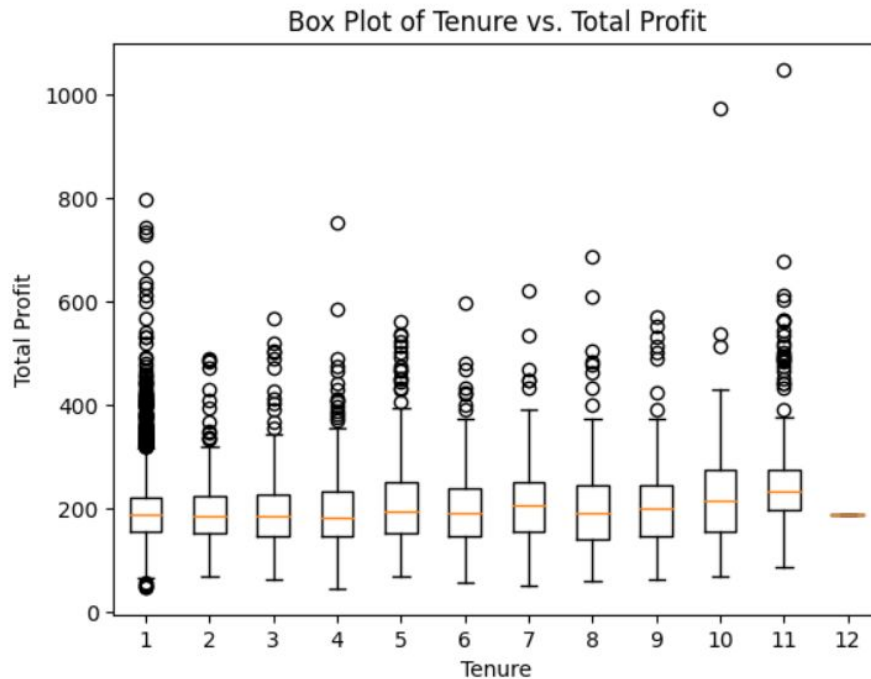


Most customers are new customers and had only made a single purchase



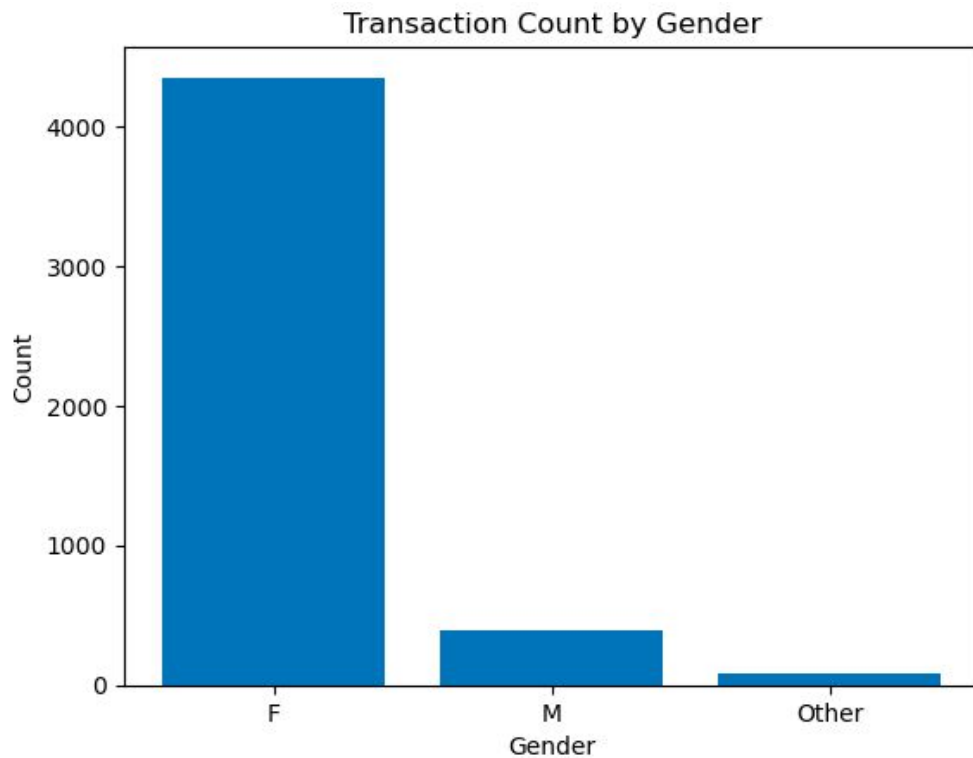
The proportion of users tends to decrease as age increases.

Visualize Customer Data



Average profit increases slightly with tenure, customers with a tenure of 1 had a more diverse range of values for profit

Visualize Customer Data



Get Product Data

	sku	Retail price	Cost	trans_count	total_sales	average_discount	trans_count_with_discount	total_returns	average_actual_price	average_profit	total_prof
0	FCS48020	79.99	48	267	352	0.273783	267	13	58.087378	10.087378	3504.4
1	FCS48021	56.99	34	3769	4127	0.084890	1581	217	52.151276	18.151276	72074.0
2	FCS48022	94.99	60	1242	1336	0.084098	510	79	87.000668	27.000668	34781.3
3	FCS48023	99.99	65	1209	1326	0.089123	532	75	91.077676	26.077676	32991.7
4	FCS48024	59.99	34	1686	1858	0.090421	744	68	54.564733	20.564733	36724.4
5	FCS48025	74.99	45	384	492	0.248568	384	25	56.347422	11.347422	5545.0
6	FCS48026	56.99	35	1651	1783	0.078619	646	98	52.508716	17.508716	30162.2
7	FCS48027	94.99	57	1238	1384	0.089701	544	65	86.468393	29.468393	38717.6
8	FCS48028	69.99	42	3369	3672	0.082012	1384	212	64.249127	22.249127	78702.7
9	FCS48029	44.99	20	4000	4364	0.083838	1671	228	41.217312	21.217312	90271.1
10	FCS48030	74.99	45	428	530	0.249065	428	24	56.310093	11.310093	5964.7

CLV Analysis

Condition: do **not** have information about churn rate of customers.

Strategy: make an assumption of churn rate **(50%/Year)**, the average life span is **02 years** (the median of tenure from our dataset) and discount rate for future cash flow is **10%**.

Since data is from first half of 2017, we annualize the cohort data (purchase_quantity, returns, spending, profit and number of order).

- Profit from sales = Revenue - cost of goods
- Return cost = 21 per return
- Order processing cost = 5 per order
- Acquisition cost: 94 (I assume it keep increasing 6 dollars as two last years)

Formula:

$M = \text{Profit from Sales} - \text{Return Cost} - \text{Order Cost}$

$CLV = M + M \cdot (1 - 0.5) / (1 + 0.1) - AC$



CLV Analysis Cont.

Conclusion: The cohort 5 with age ranging from 32 to 34 has highest CLV of **469**. For new customer acquisition campaign, we should first target on 32-34 years old customer.

	average_price	average_discount	purchased_quantity	product_return	spend	profit	number_order	CLV
cohort_group								
1	62.612248	0.098126	22.428571	1.194617	1411.883375	409.263188	16.169772	347.203448
2	62.998078	0.098139	22.784679	1.186335	1446.618944	418.354555	16.525880	358.090344
3	62.274332	0.090217	18.837306	1.036269	1182.893492	351.447275	13.262176	289.090525
4	66.592105	0.082701	18.424870	0.853886	1251.832332	379.820518	13.127461	306.911427
5	73.386582	0.084275	22.974093	1.255959	1686.431399	497.974197	16.918135	468.921300



RFM Analysis

Counts of different RFM scores

111	1159
222	1157
122	805
211	804
212	231
121	230
221	221
112	220

We have two significant groups (based on recency, frequency, and monetary behavior.), each containing a substantial number of customers.

- 1. Group 1 - RFM Score 111:**
 - a.** Low recency, infrequent purchases, and low spending.
 - b.** Action: Targeted promotions to re-engage and win back their business.
- 2. Group 2 - RFM Score 222:**
 - a.** High recency, frequent purchases, and higher spending.
 - b.** Action: Maintain excellent service, explore upselling opportunities, and offer personalized loyalty benefits.



RFM Analysis Cont.

customer_id	id	age	gender	tenure	last_order_date	number_order	total_spend	day_from_last_order	recency_score	frequency_score	monetary_score	rfm_score
1168	129442	33	F	10	2017-07-20	32	3303.55	0 days	2	2	2	222
4411	182944	25	M	3	2017-07-08	31	2844.59	12 days	2	2	2	222
608	119839	33	F	9	2017-07-05	31	2825.61	15 days	2	2	2	222
4451	183631	24	M	0	2017-07-20	28	2798.62	0 days	2	2	2	222
3196	163149	21	F	0	2017-07-08	29	2778.56	12 days	2	2	2	222
...
3554	168837	30	F	5	2017-03-20	3	186.97	122 days	1	1	1	111
1722	138956	28	F	6	2017-03-11	3	186.97	131 days	1	1	1	111
281	114822	28	F	0	2017-05-08	3	186.97	73 days	1	1	1	111
4779	189125	28	Other	3	2017-05-17	3	165.97	64 days	1	1	1	111
2640	153895	30	F	0	2017-07-03	3	161.97	17 days	1	1	1	111



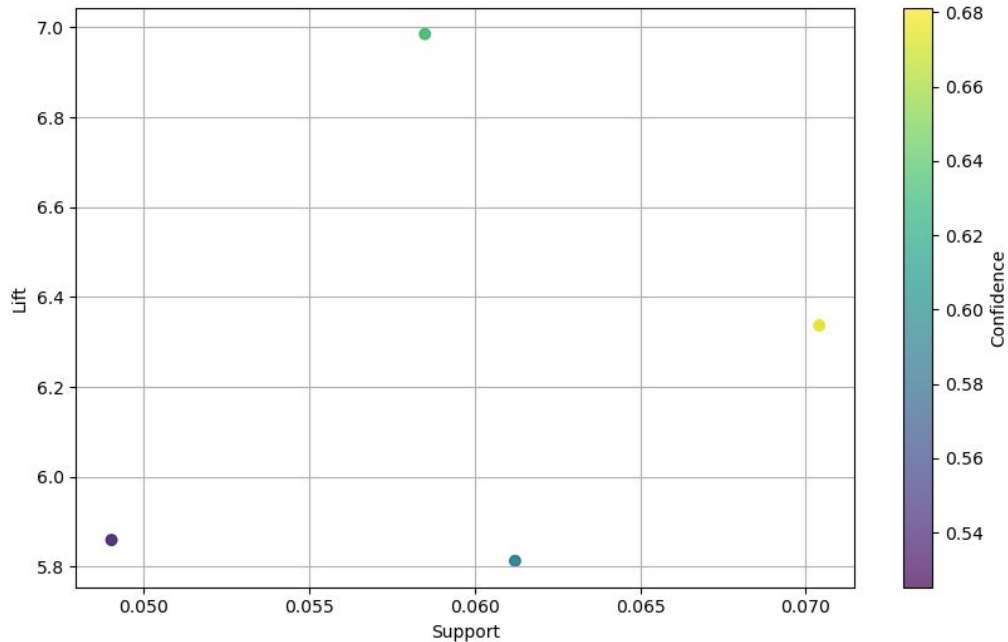


Marketing Basket Analysis (MBA)

A lift of 6.985086 for the SCS72048 (antecedent) and FCS48028 (consequent) pair suggests these items are nearly seven times more likely to be bought together than independently, indicating a strong purchase association.

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
2	(SCS72048)	(FCS48028)	0.089599	0.093459	0.058492	0.652821	6.985086	0.050118	2.611161	0.941165
3	(FCS48028)	(SCS72048)	0.093459	0.089599	0.058492	0.625854	6.985086	0.050118	2.433281	0.945173
6	(SCS72043)	(SCS72044)	0.107513	0.103347	0.070407	0.654870	6.336624	0.059296	2.598013	0.943641
7	(SCS72044)	(SCS72043)	0.103347	0.107513	0.070407	0.681268	6.336624	0.059296	2.800122	0.939256
5	(FCS48031)	(SCS72042)	0.093320	0.089654	0.049021	0.525298	5.859151	0.040654	1.917719	0.914686
4	(SCS72042)	(FCS48031)	0.089654	0.093320	0.049021	0.546778	5.859151	0.040654	2.000521	0.911002
0	(SCS72046)	(FCS48021)	0.100680	0.104597	0.061214	0.608000	5.812809	0.050683	2.284192	0.920658
1	(FCS48021)	(SCS72046)	0.104597	0.100680	0.061214	0.585236	5.812809	0.050683	2.168270	0.924685

Marketing Basket Analysis (MBA) Result



	bundle	retail_price	cost	margin
0	SCS72046 + FCS48021	156.98	94	62.98
1	SCS72048 + FCS48028	139.98	84	55.98
2	FCS48031 + SCS72042	83.98	57	26.98
3	SCS72043 + SCS72044	119.98	70	49.98

Clustering analysis



	Cluster 0	Cluster 1
transaction_id	25076.346	22601.290
price	98.032	57.373
quantity	1.098	1.101
age	29.802	26.805
tenure	5.055	2.945
return	0.058	0.057
Retail price	98.032	57.373
Cost	60.424	33.948
actual_price	89.730	51.804
profit	29.306	17.856
total_profit	30.814	18.901
spend	107.455	63.346
sku_FCS48020	-0.000	0.007
sku_FCS48021	-0.000	0.104
sku_FCS48022	0.125	0.000
sku_FCS48023	0.122	0.000
sku_FCS48024	-0.000	0.046
sku_FCS48025	0.000	0.010
sku_FCS48026	-0.000	0.045
sku_FCS48027	0.125	0.000
sku_FCS48028	0.001	0.093
sku_FCS48029	0.000	0.110
sku_FCS48030	0.001	0.012
sku_FCS48031	0.000	0.093
sku_FCS48032	-0.000	0.047
sku_FCS48033	0.125	0.000
sku_SCS72041	0.000	0.016
sku_SCS72042	-0.000	0.092
sku_SCS72043	-0.000	0.107
sku_SCS72044	-0.000	0.106
sku_SCS72045	0.122	0.000
sku_SCS72046	0.378	0.000
sku_SCS72047	0.000	0.012
sku_SCS72048	0.001	0.092
sku_SCS72049	0.000	0.008
discount_0.0	0.579	0.550
discount_0.1	0.082	0.075
discount_0.15	0.084	0.079
discount_0.2	0.083	0.092
discount_0.25	0.084	0.104
discount_0.3	0.087	0.100
gender_F	0.792	0.903
gender_M	0.194	0.081
gender_Other	0.014	0.017

The **price**, **actual_price**, and **Spend** are higher for Cluster 0, meaning that customers in Cluster 0 tend to purchase more **expensive** items than those in Cluster 1

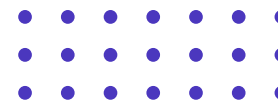
The **tenure** for Cluster 0 (older) is higher, indicating that customers in this cluster have been with the company for a **longer time** compared to Cluster 1.

Product **sku_SCS72046** is most popular product among older customer cluster, occupy 37.8% of associated transactions.

Cluster1 (**younger** customers) would be **more sensitive** to **deep** discount.

03

Recommendations





Approach 1: Targeting Current Customers

Goal 1: Retaining Churned Customers and Increasing Spending

Strategy:

Target Customers: RFM Scores 112 (profitable customers but did not purchase frequently for a while) **and 122** (profitable and high frequent buyers but already left us for a while).

Offer: Personalized Coupons or Discounts for Routine Products or Bundles

Objective:

Entice customers to make a purchase

Timeline: Send out promotions every 4 months (Get 20% when you spend \$100+ if you are a member)



Approach 1: Targeting Current Customers

Goal 2: Up-selling and Cross-selling

Strategy:

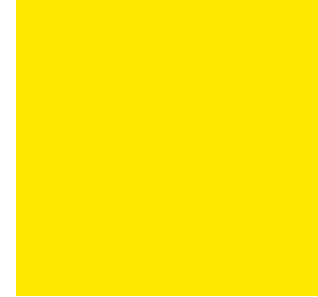
Target Customers: Lower "Total Spending" and "Average Price" compared to others

Approach: Market Basket Analysis (MBA) and Recommender System

Objective:

Suggest alternative luxury products/bundles and encourage trials

Timeline: Send out promotions once a month for Buy 2 Products above \$50 Get 1 Free.



Approach 2: New Customer Acquisition Strategy

Who: CLV Analysis suggests: 32-34 Years Old Age Group

How: Offer Promotions for Popular and Basic Bundles

What Bundles? (Top 4 from MBA analysis)

SCS72048 & FCS48028 ; SCS72044 & SCS72043 ; **SCS72046** & FCS48021; FCS48031 & SCS72042

↑
Popular in older group

Who:: around Years young Age Group

How: Offer deep Promotions Ex: 25%, 30% (Clustering Analysis)

What Bundles? (Top 3 from Clustering analysis)

SCS72044 & SCS72043; SCS72046 & FCS48021

Timeline: Throughout the year: Sign up as a member, get 10% off and a free gift!

THANK YOU!