



CUSTOMER ANALYSIS PROJECT

**WANG YUHUI
YANG YINYU**

**CHEN HONGYANG
ILANSURYA**

AGENDA

01

INTRODUCTION

02

EDA ANALYSIS

03

ANALYSIS STRUCTURE

04

RFM ANALYSIS

05

GENERAL MBA ANALYSIS

06

PERSONALIZED RECOMMENDATION

07

CHANNEL ANALYSIS

08

INSIGHTS & RECOMMENDATIONS

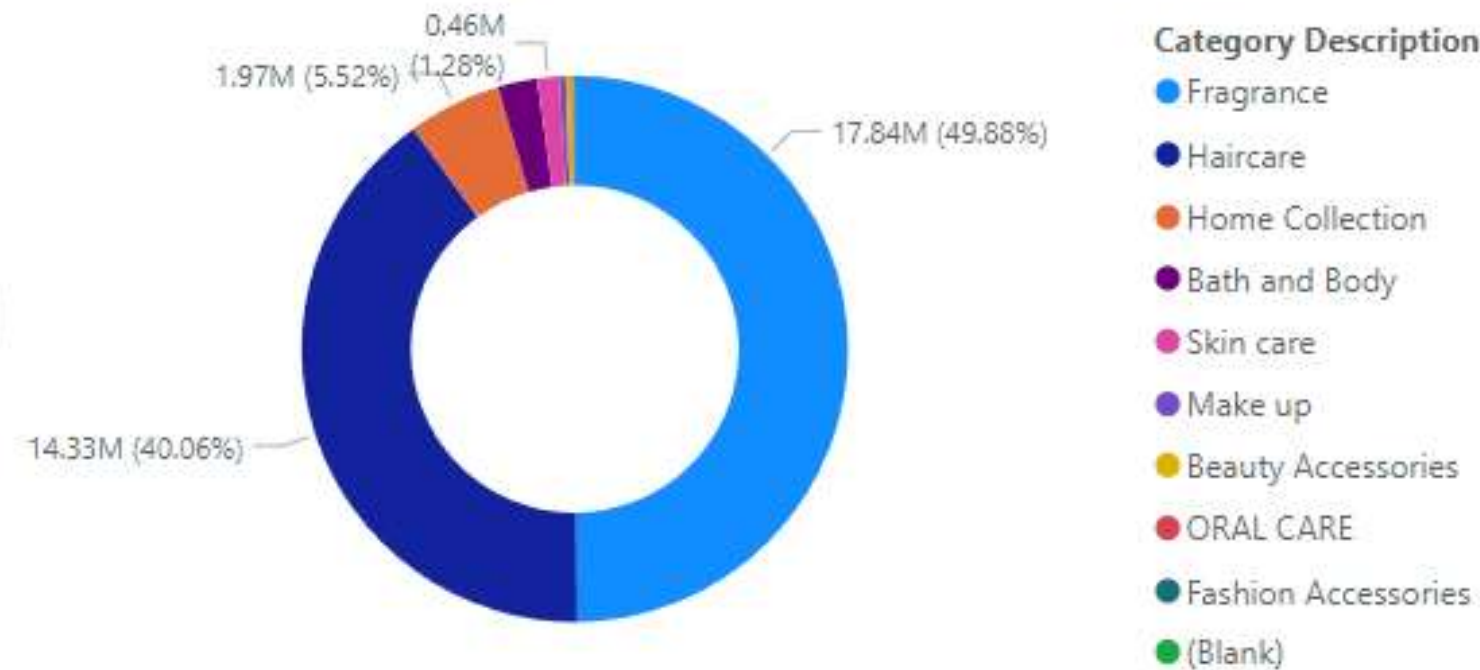
The background features abstract line art. On the left, a series of curved lines form a funnel-like shape that tapers towards the top. On the right, several vertical, wavy lines flow downwards. The central area is left blank, providing space for the title and list.

INTRODUCTION

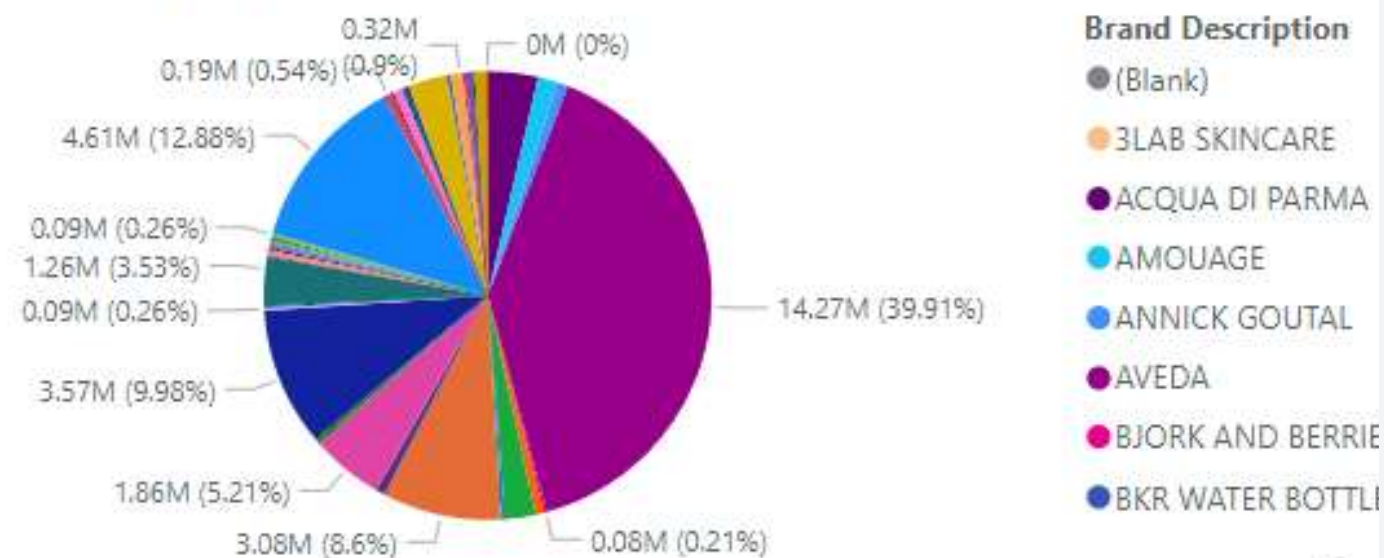
- Quantify customer segment value and accurately predict customer churn using prediction and RFM analysis
- Identify highly correlated product bundles through market basket analysis
- Personalize product promotion strategies for each customer and implement a data-driven approach to increase customer satisfaction and loyalty in the competitive essentials market
- Customize product recommendation strategies based on different shopping channels to increase revenue while improving customer satisfaction

EDA ANALYSIS

Revenue by Category



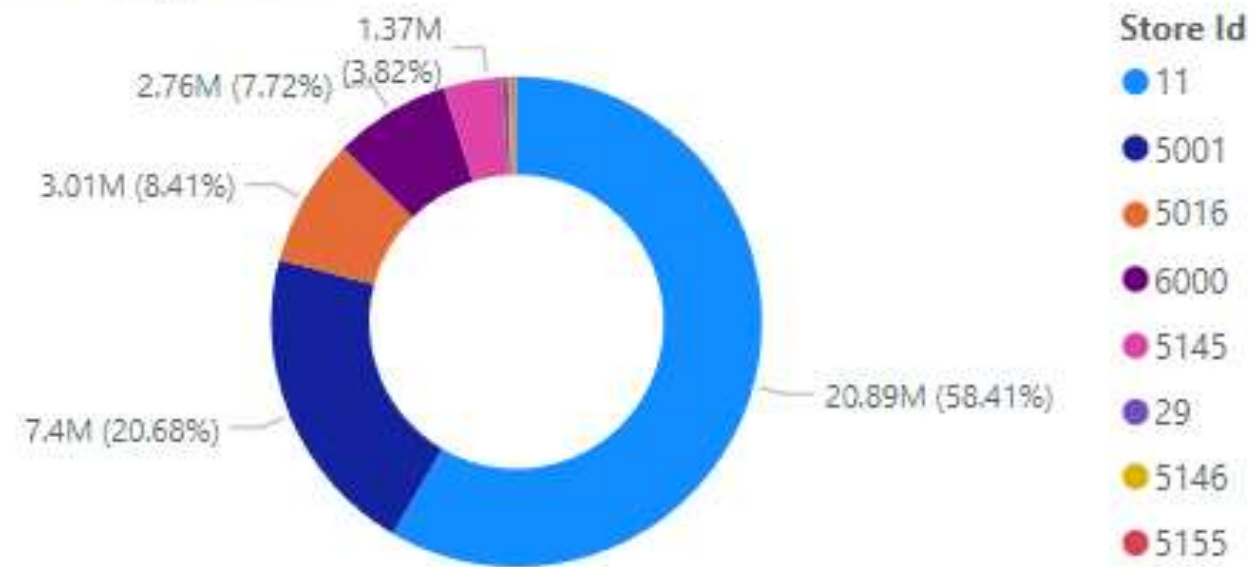
Revenue by Brand



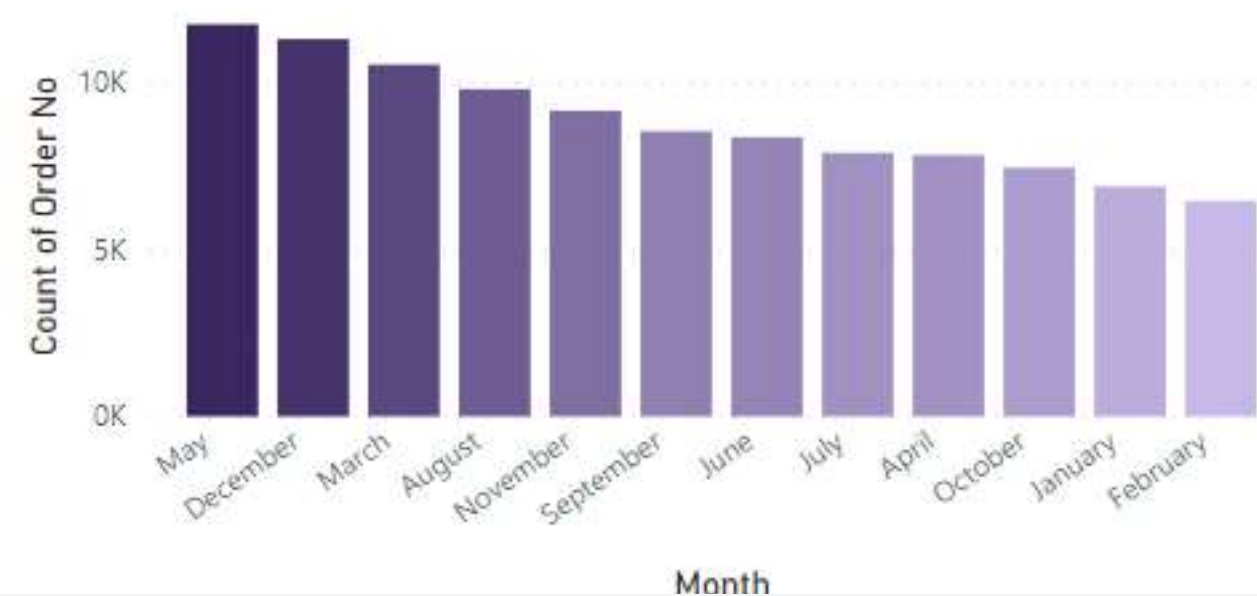
- Majority of the revenue is generated from fragrance products followed by haircare.
- 40% of the revenue is generated from the brand called Aveda
- Maison Francis brand receives the greatest number of orders

EDA ANALYSIS

Revenue by Store



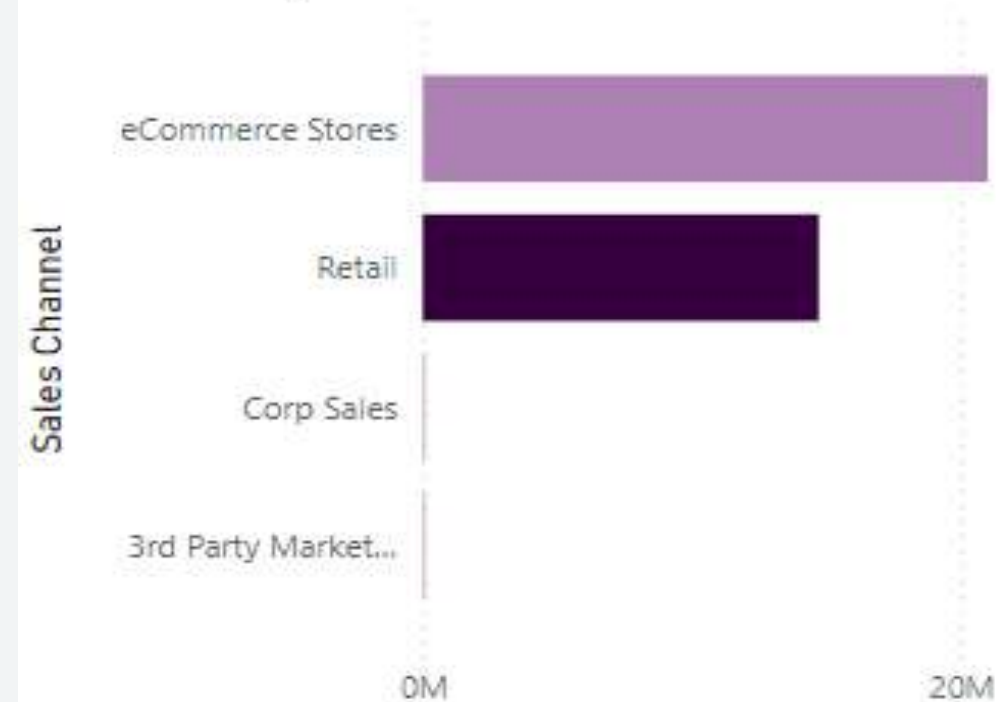
Count of Order No by Month



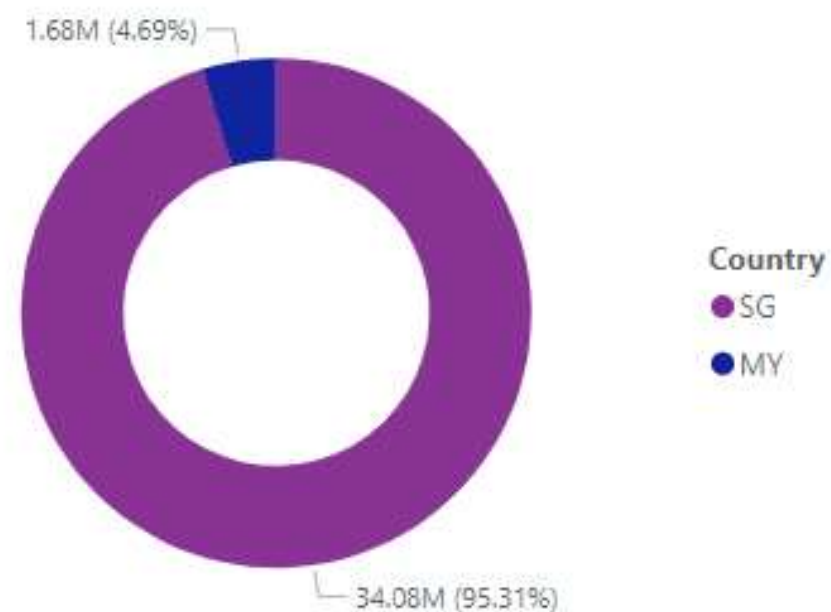
- High value customers place their order in store 11
- The sale of products tend to peak during May and December

EDA ANALYSIS

Revenue by sales channel

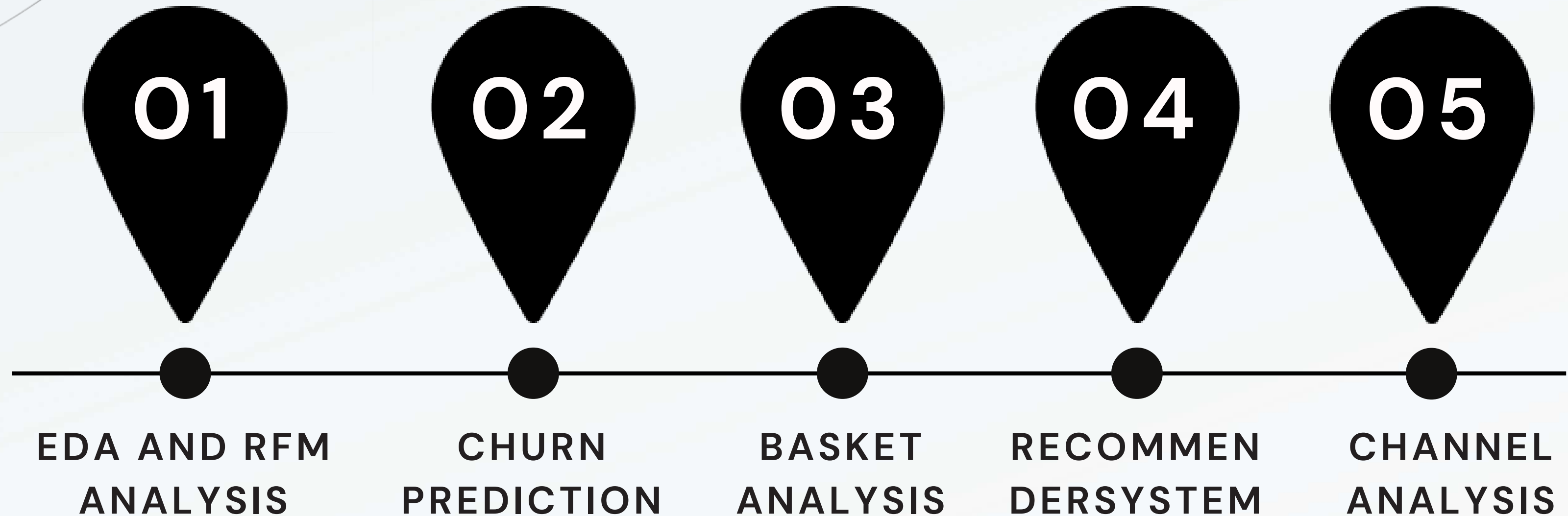


Revenue by Country



- E-commerce contributes to 58% of the revenue generated in Singapore whereas retail sales contributes more to Malaysia
- Majority of the fragrance products are sold in retail stores whereas hair care products are mostly sold in e-commerce stores
- Majority of the revenue is generated from 20% of the customers

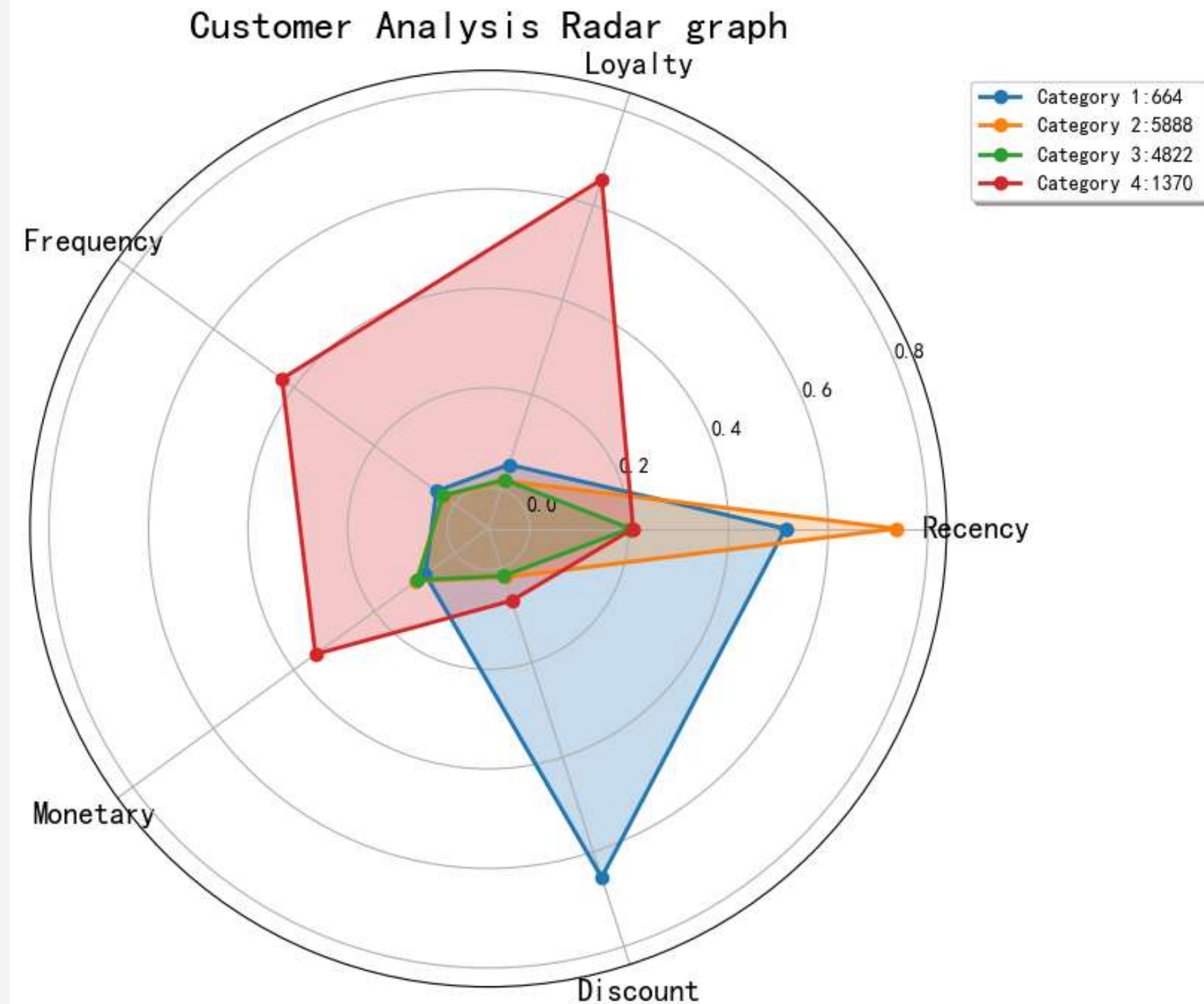
ANALYSIS STRUCTURE



RFM & K-MEANS

What we will do for RFM analysis is to

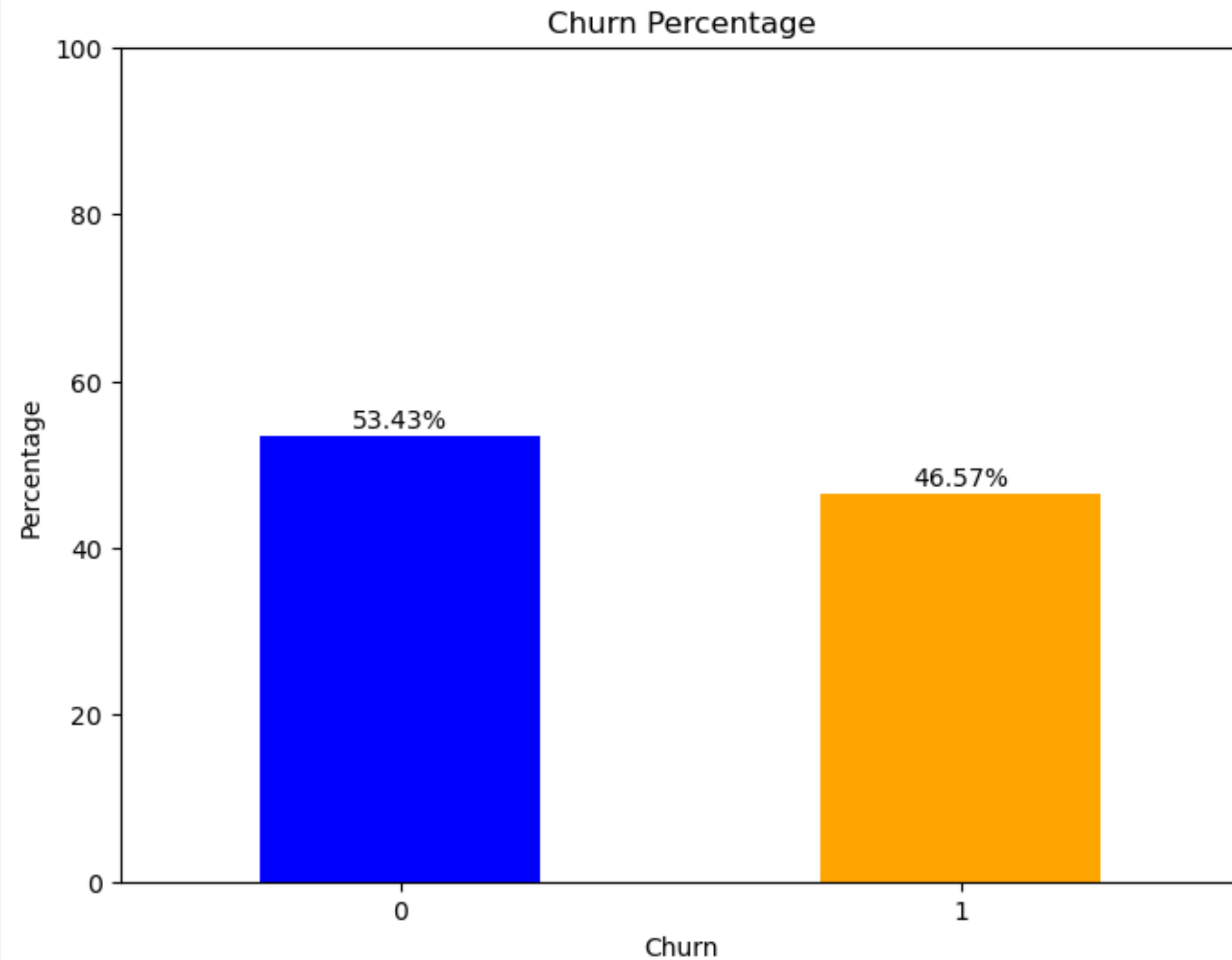
- Aggregate Recency, Frequency, Monetary and some other metrics
- Eliminate the outliers and then normalize the data
- Use K-Means to segment the customers



CHURN PREDICTION

Churn Definition and data subset

- Customer is considered Churn is defined as if a customer did not purchase over 3 months
- Only Category 1 and 4 customer segmentation is considered
- Only customers purchase more than once is considered

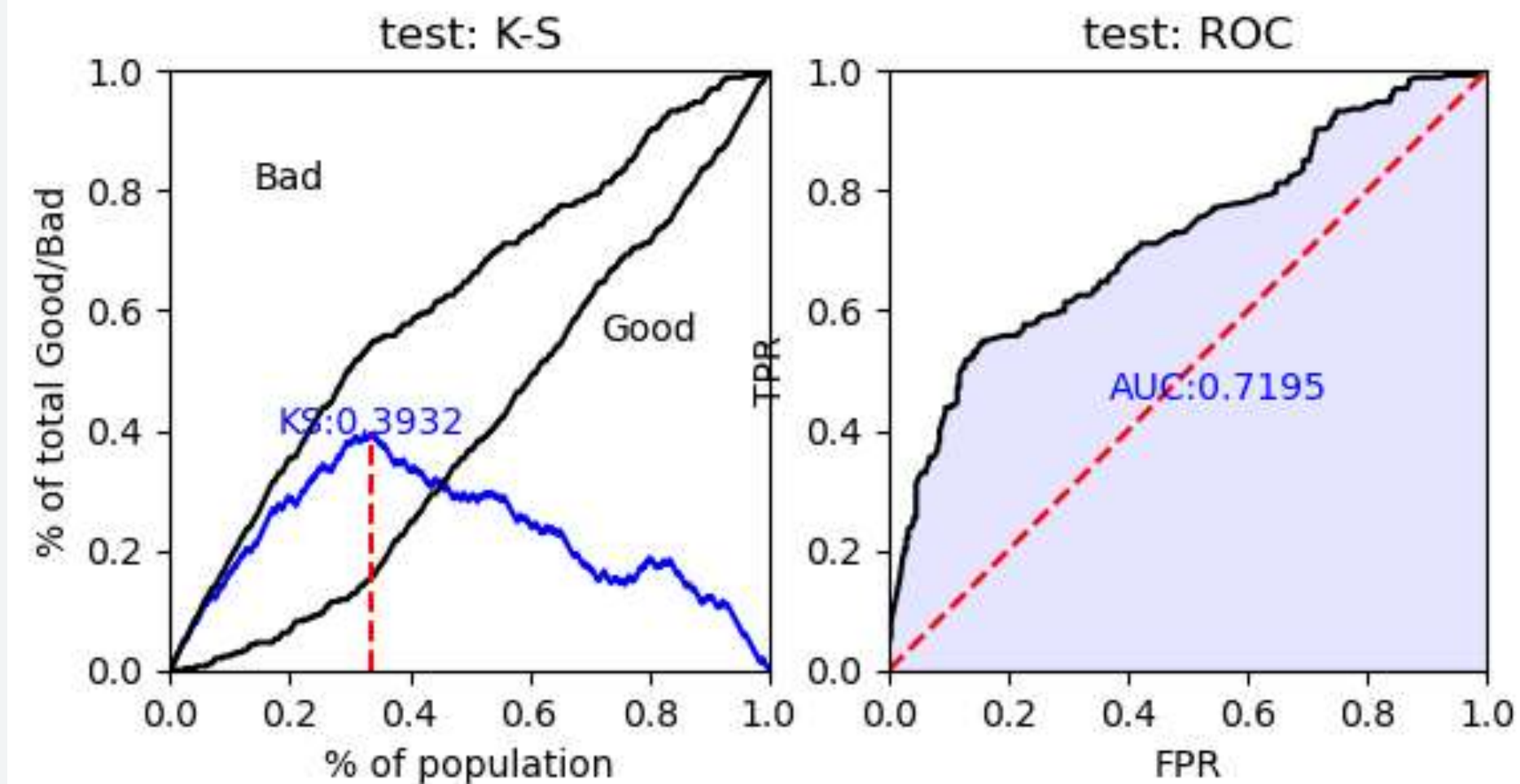


CHURN PREDICTION-XGBOOST

XGBoost

Summary report for XGB Boost
Classification Report:

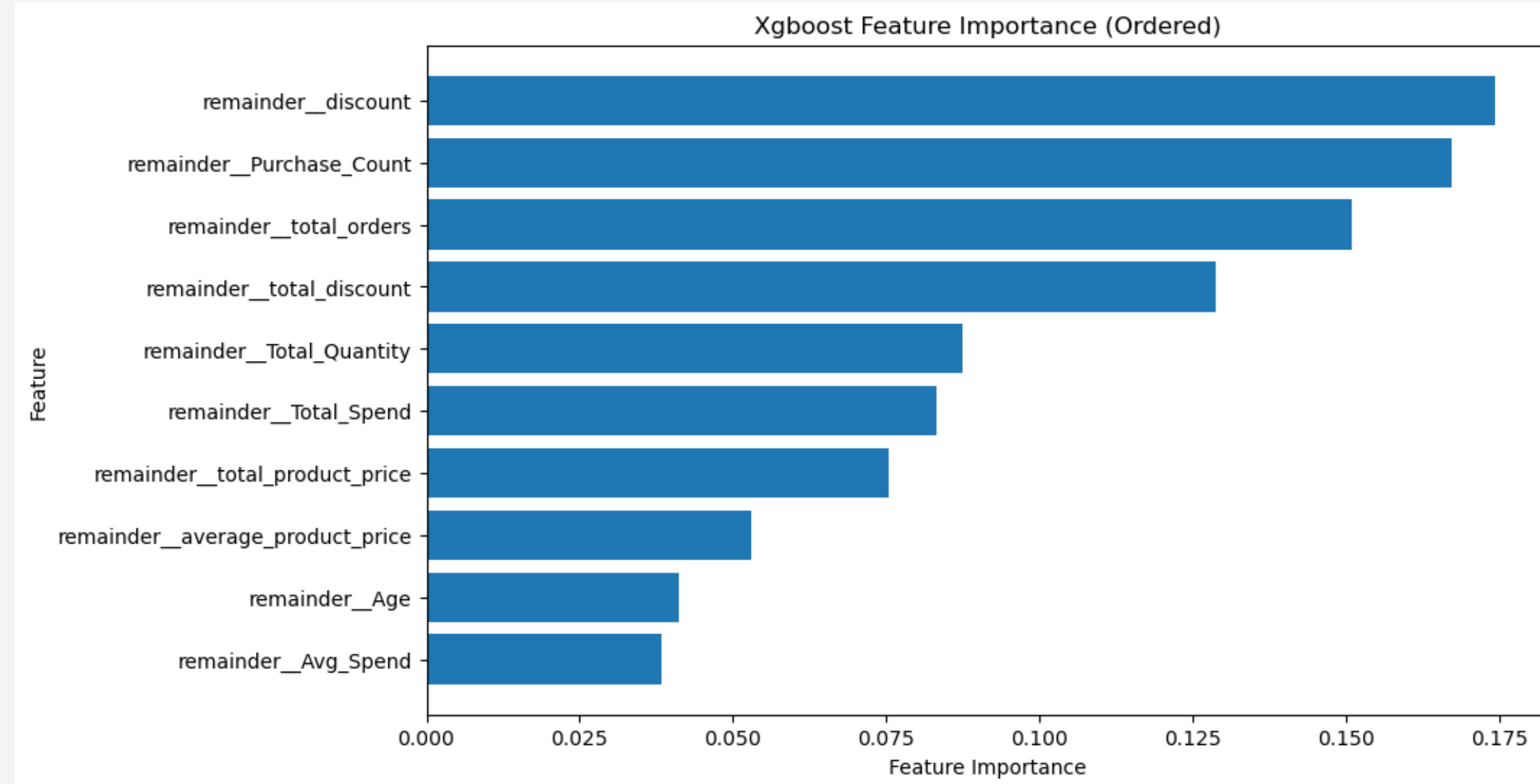
	precision	recall	f1-score	support
0	0.67	0.87	0.76	306
1	0.77	0.52	0.62	267
accuracy			0.70	573
macro avg	0.72	0.69	0.69	573
weighted avg	0.72	0.70	0.69	573



CHURN PREDICTION-XGBOOST

XGBoost

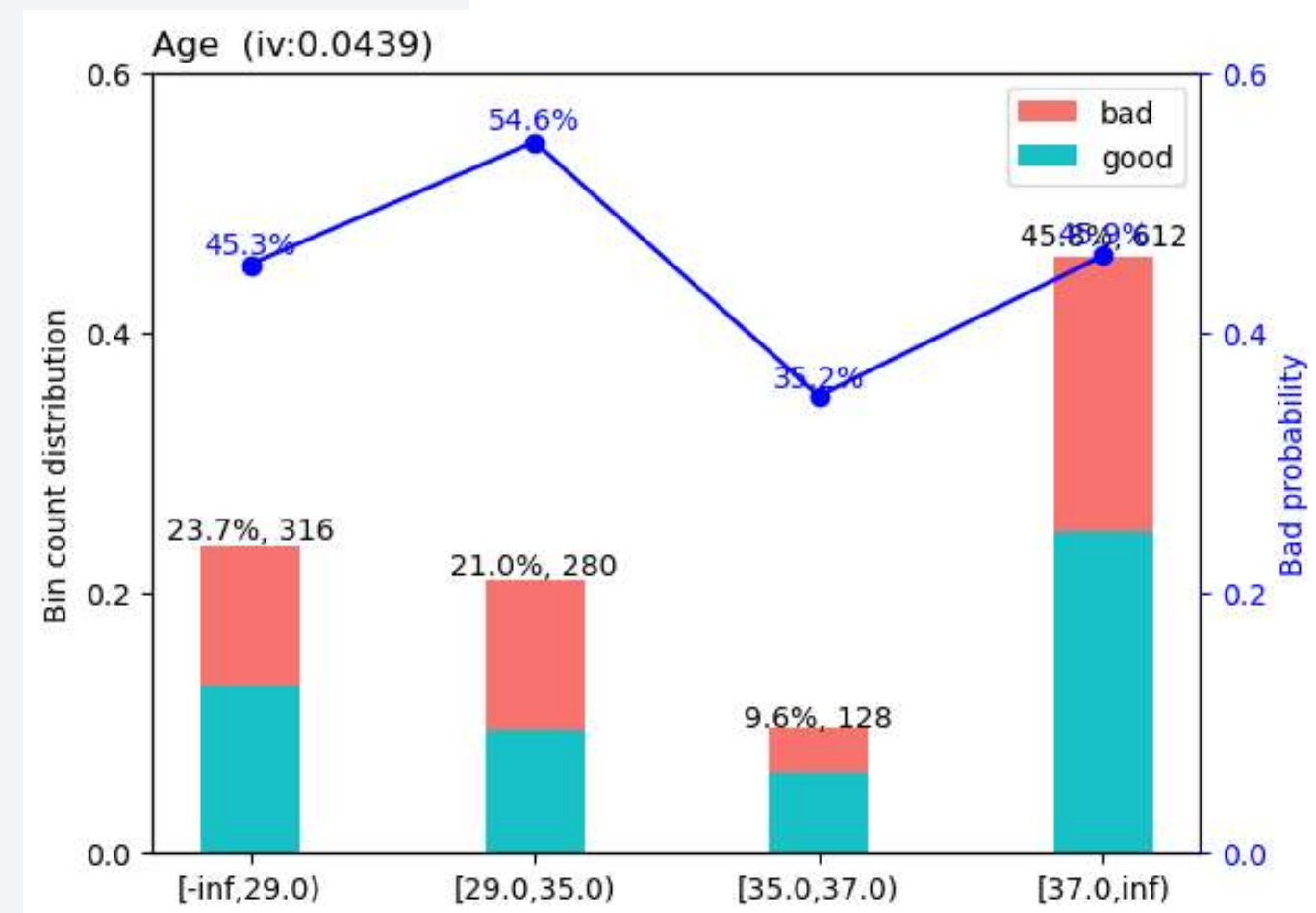
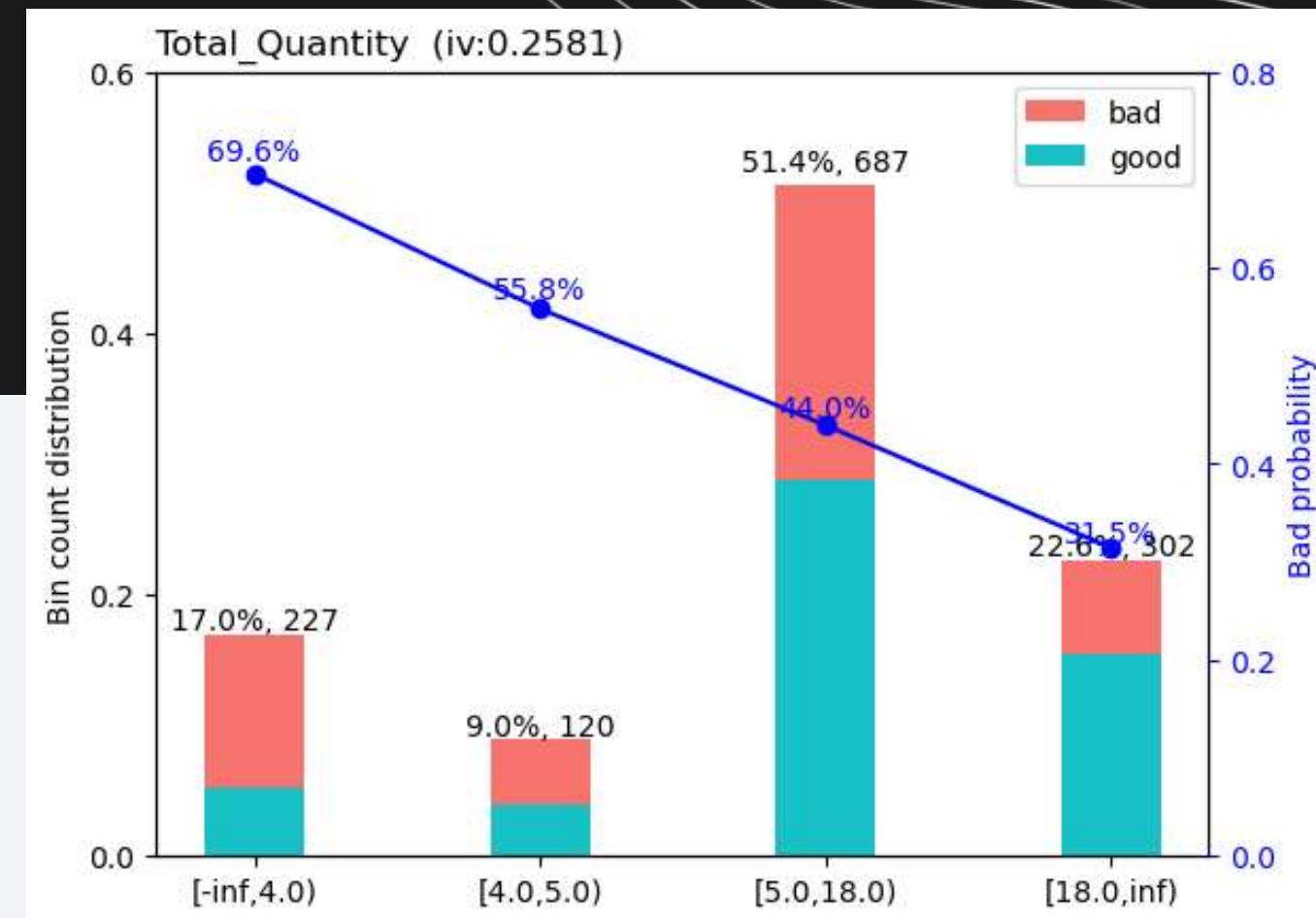
Most important aggregated variable is
discount and total number items
customer purchase



CHURN PREDICTION- LOGISTIC REGRESSION

LOGISTIC REGRESSION

- Bin the variables
- Build the scorecard
- Give more interpretability



CHURN PREDICTION-SCORECARD

SCORECARD

- Scorecard provide much better interpretation and business insights

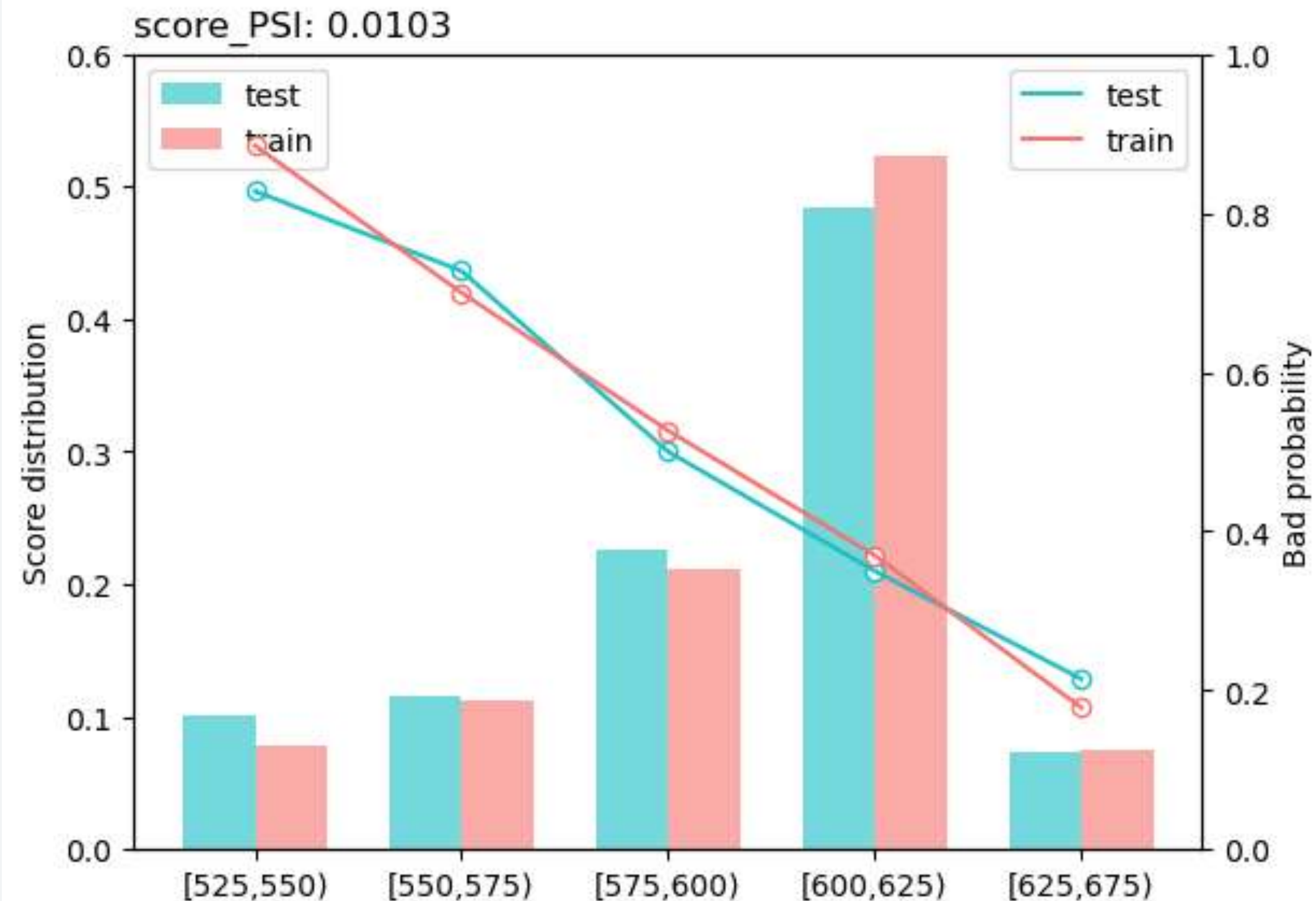
	variable	bin	points
0	basepoints	NaN	593.0
0	average_product_price	[-inf,80.0)	-0.0
1	average_product_price	[80.0,110.0)	0.0
2	average_product_price	[110.0,120.0)	2.0
3	average_product_price	[120.0,inf)	-0.0
4	Total_Quantity	[-inf,4.0)	-5.0
5	Total_Quantity	[4.0,5.0)	-2.0
6	Total_Quantity	[5.0,18.0)	1.0
7	Total_Quantity	[18.0,inf)	3.0
11	Total_Spend	[-inf,200.0)	-14.0
12	Total_Spend	[200.0,400.0)	-6.0
13	Total_Spend	[400.0,3700.0)	3.0
14	Total_Spend	[3700.0,inf)	11.0
8	discount	[-inf,0.09)	0.0
9	discount	[0.09,0.16)	-1.0
10	discount	[0.16,inf)	-0.0
15	Purchase_Count	[-inf,4.0)	-4.0
16	Purchase_Count	[4.0,13.0)	-0.0
17	Purchase_Count	[13.0,35.0)	2.0
18	Purchase_Count	[35.0,inf)	6.0
19	Age	[-inf,29.0)	1.0
20	Age	[29.0,35.0)	-3.0
21	Age	[35.0,37.0)	5.0
22	Age	[37.0,inf)	0.0

23	total_orders	[-inf,2.0)	-25.0
24	total_orders	[2.0,3.0)	-5.0
25	total_orders	[3.0,10.0)	6.0
26	total_orders	[10.0,inf)	34.0
27	total_product_price	[-inf,400.0)	2.0
28	total_product_price	[400.0,1000.0)	0.0
29	total_product_price	[1000.0,4500.0)	-1.0
30	total_product_price	[4500.0,inf)	-2.0
31	total_discount	[-inf,5.0)	6.0
32	total_discount	[5.0,65.0)	-11.0
33	total_discount	[65.0,145.0)	-8.0
34	total_discount	[145.0,inf)	3.0

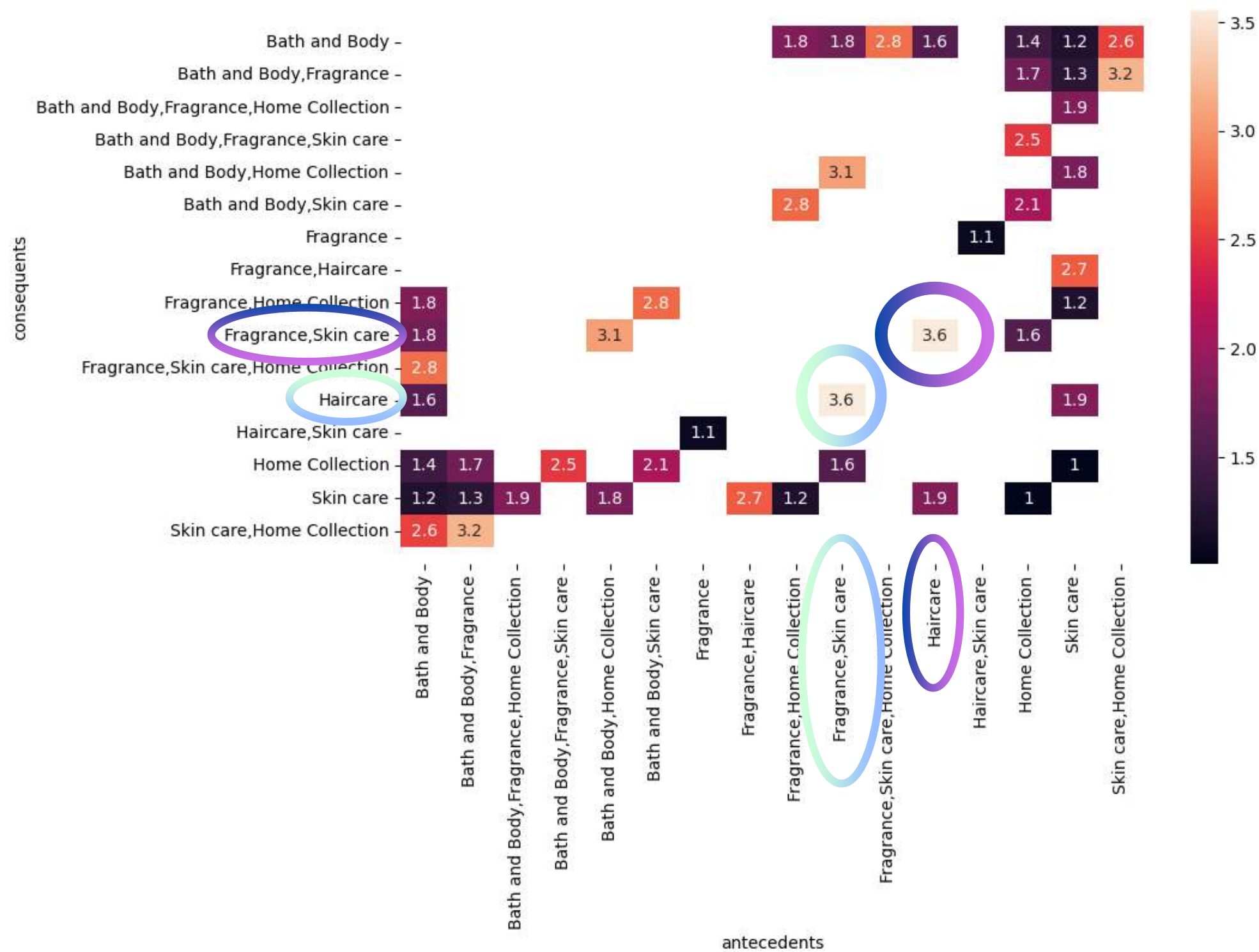
CHURN PREDICTION-SCORECARD

LOGISTIC REGRESSION

- Make strategies based on point level of the customers



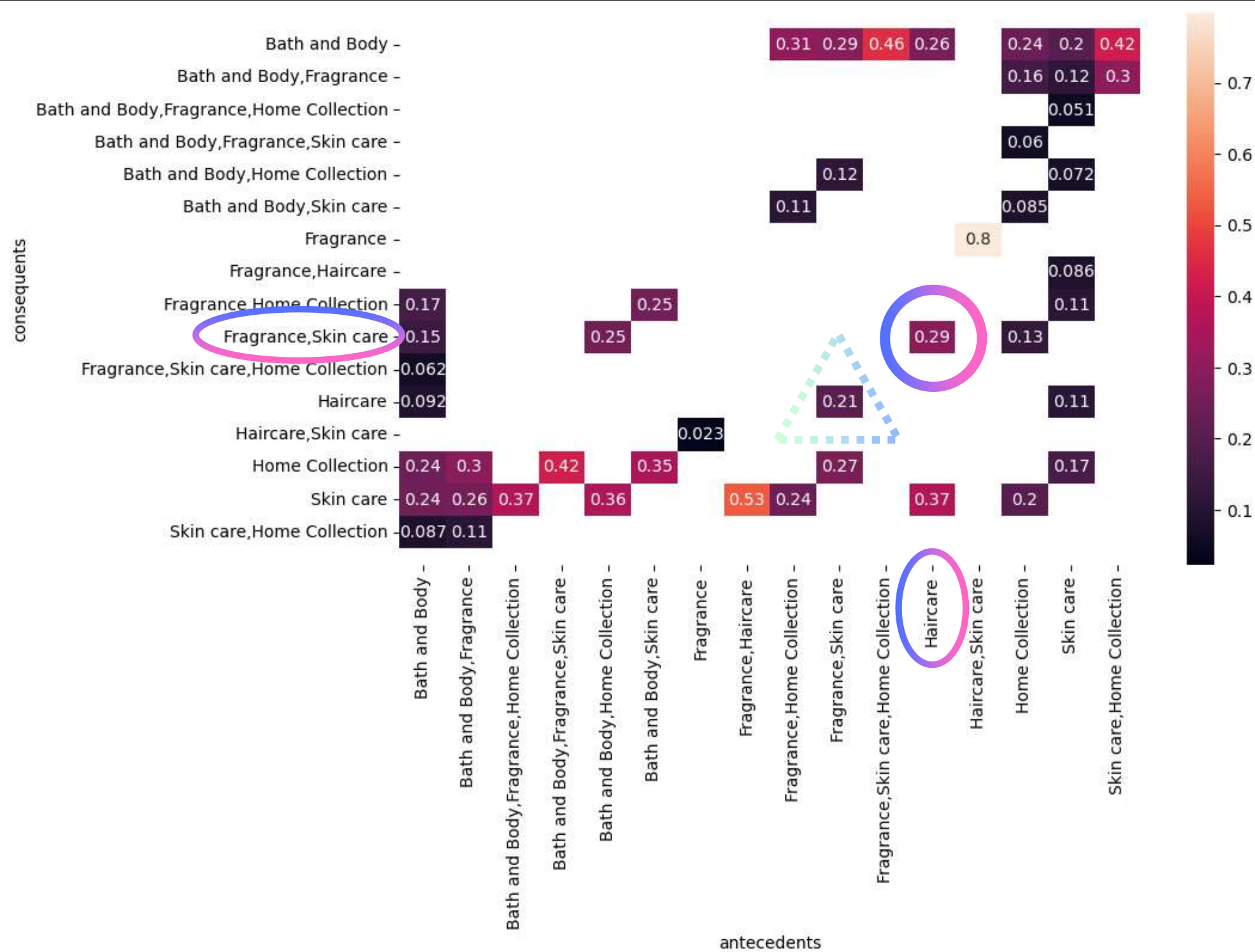
GENERAL MBA



Insights from Association Rules

- MBA result based on **Lift** value
- The highest Lift value is **3.6**
- Result shows customers who buy ***Fragrance*** and ***Skin care*** products are significantly more likely to also purchase ***Haircare*** products together

GENERAL MBA



- MBA result based on **Confidence** value
- Confidence value of rule **{Fragrance, Skin care} => {Haircare}** is **0.21** while Confidence value of rule **{Haircare} => {Fragrance, Skin care}** is **0.29**
- Suggesting that customers who buy **Haircare** products are more likely to also purchase **Fragrance and Skin care** products than vice versa

RECOMMENDER SYSTEMS

RMSE: 0.0535

RMSE: 0.05347421166015877

Top 10 recommendations for user 2218642:

	Article Code	Estimated Rating
0	AWLE010000	0.057260
1	AR5Y010000	0.057199
2	MFK1CMH008	0.056683
3	BKFN4NP010000	0.053666
4	MFK202292401	0.052538
5	MLF192	0.052042
6	AR5L010000	0.048727
7	NG35-1	0.048088
8	ADP067	0.047075
9	TRST3M026	0.046325

Collaborative Filtering

- To find top 10 product recommendation for every user
- Offer discounts to the particular customers for these products
- Offer these products as bundles
- Increased revenue generation

RECOMMENDER SYSTEMS

Article Code

XM23B190A 0.826123

XM23B70A 0.818886

XM22VSP 0.717079

XM23COF70 0.682898

XM23LANTERN 0.680308

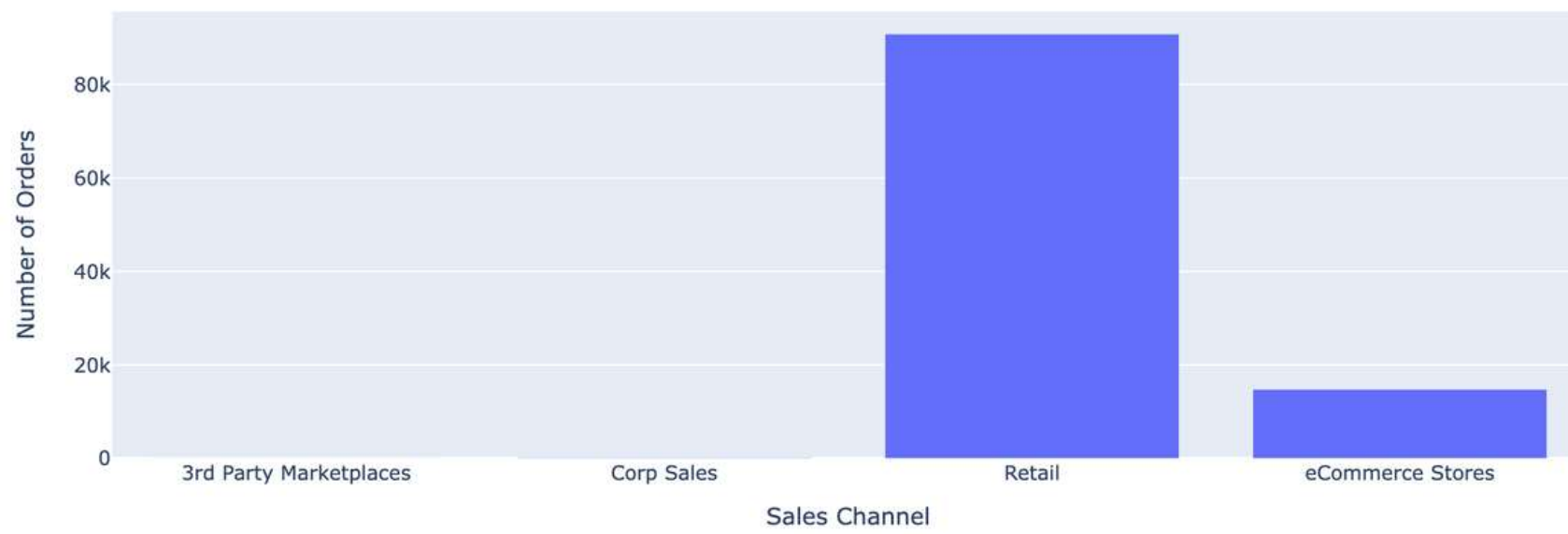
Name: XM23VSP, dtype: float64

Cosine Similarity

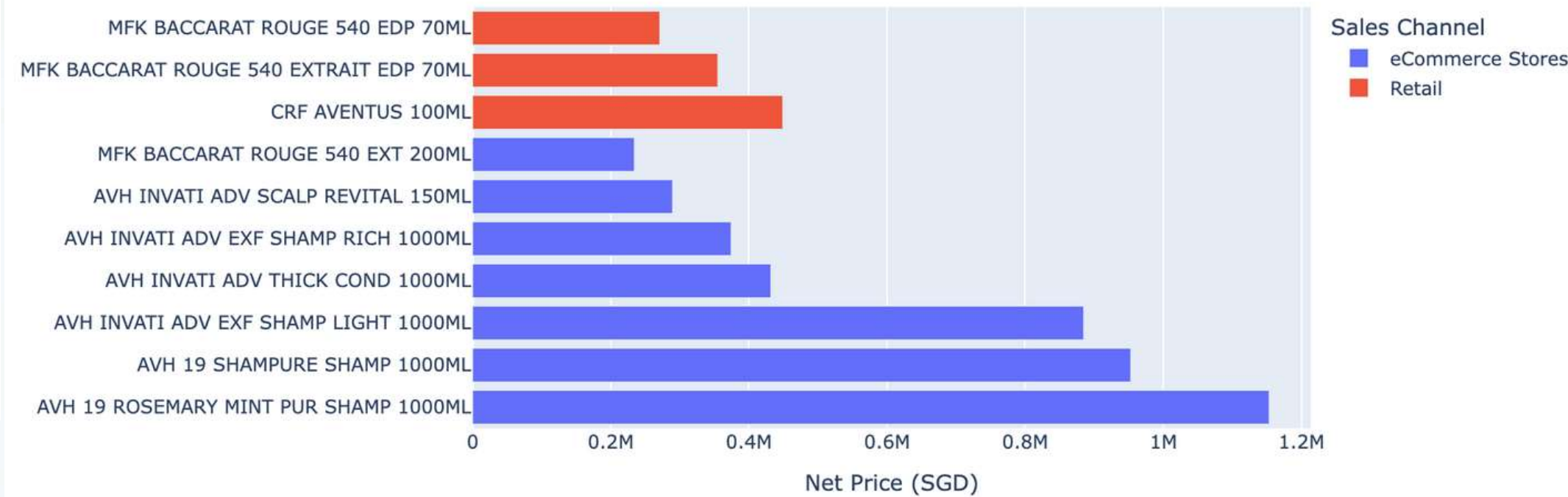
- Similar product recommendations
- These products can be placed together in a in stores
- In online stores recommended these as similar products

CHANNEL ANALYSIS

Number of Orders by Sales Channel



Top-Selling Products by Sales Channel



Total Sales by Sales Channel

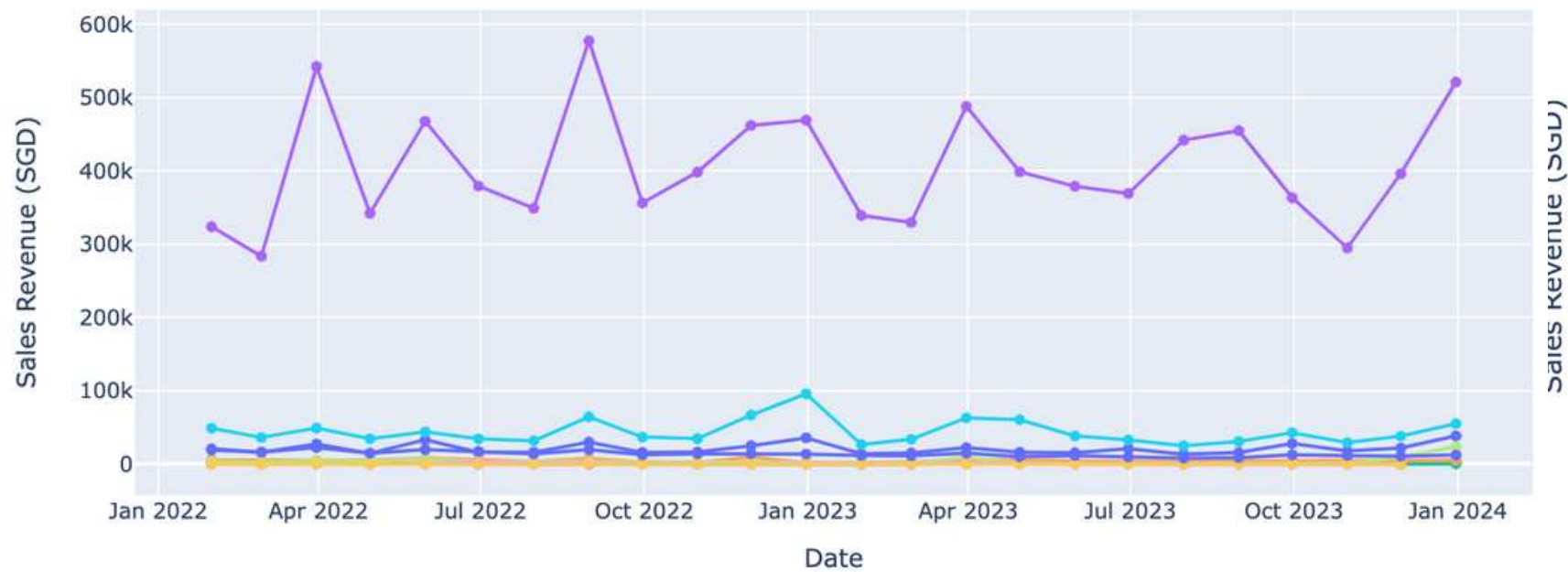


- The two most important channels with highest revenue are retail and e-commerce store.
- Products with large content are sold more via e-commerce.

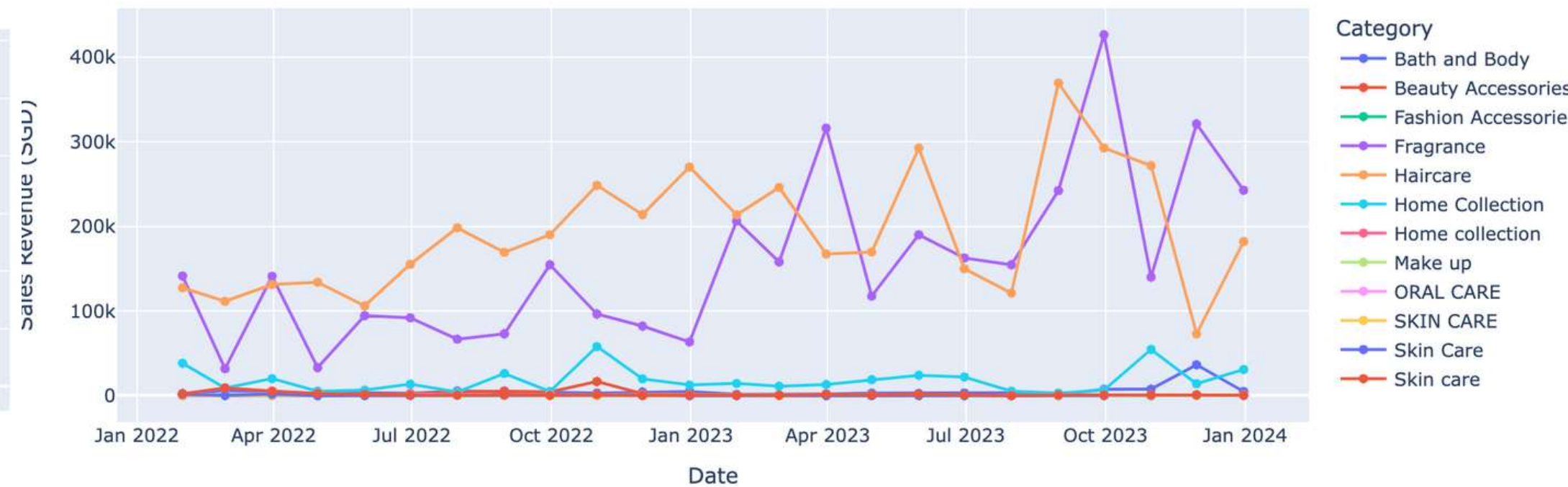
CHANNEL ANALYSIS

Monthly Revenue by Channels

Monthly Sales Revenue for Each Category - Retail Channel



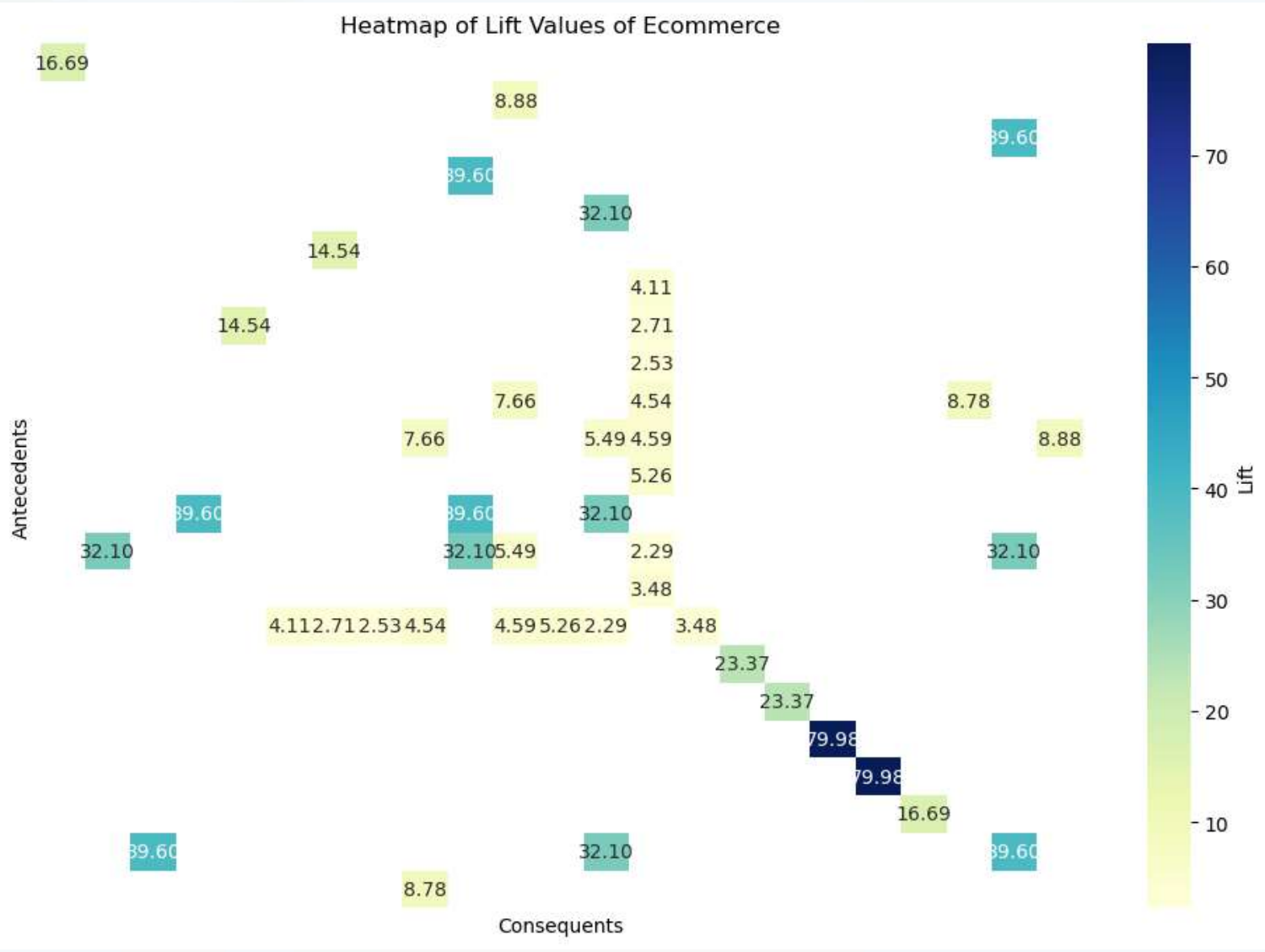
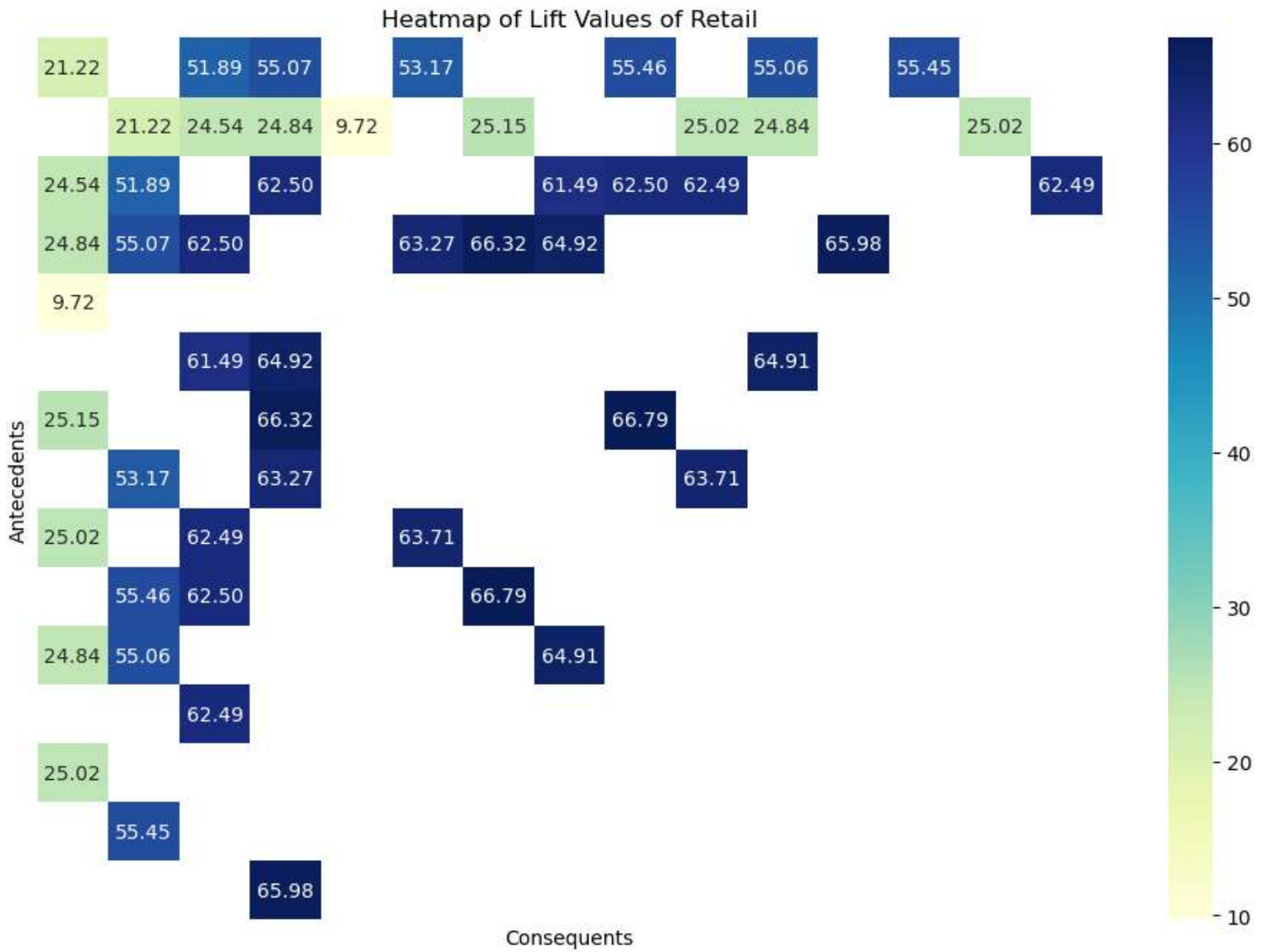
Monthly Sales Revenue for Each Category - eCommerce Stores Channel



From two plots above, we can see that the highest revenue of categories for each channel is Fragrance and Haircare. So, we do a product recommendation system based on these two categories.

CHANNEL ANALYSIS

Lift heatmap of two channels



	Item	Category Description
0	MVF NEROLI NASIMBA PERFUME 1,5ML SAMPLE	Fragrance
1	CRF AVENTUS 2ML VIAL 1=100	Fragrance
2	BKF CAN'T STOP LOVNG YOU VOC 1.5ML	Fragrance
3	LWF CARTERITA L001 EDP 2MLX2 VIAL 1=24	Fragrance

	Item	Category Description
0	AVH NUTRIPLENISH COND DEEP 1000ML	Haircare
1	AVH COLOR CONSERVE SHMP 1000ML	Haircare
2	AVH 19 SHAMPURE CONDITIONER 250ML	Haircare
3	AVH 19 SHAMPURE SHAMP 1000ML	Haircare
4	AVH NUTRIPLENISH COND LT 1000ML	Haircare
5	AVH NUTRIPLENISH SHAMP DP 1000ML	Haircare
6	AVH INVATI ADV EXF SHAMP LIGHT 1000ML	Haircare

INSIGHTS & RECOMMENDATIONS



1. Focus on High-Revenue Products:

- Prioritize fragrance and haircare products, which drive most of the revenue and try to diversify in other categories such as body wash etc so that even if one segment sales drop others will still make profits.

2. Target Key Markets:

- Leverage the significant sales contributions from Singapore by enhancing e-commerce experiences, especially for larger products.
- Boost retail efforts in Malaysia, where physical store sales are stronger.

3. Capitalize on High-Value Customers:

- Develop targeted strategies for high-value customers concentrated in Store 11. This includes personalized retention strategies using RFM (Recency, Frequency, Monetary) analysis.

INSIGHTS & RECOMMENDATIONS



4. Seasonal Campaigns:

- Plan major promotional campaigns during peak sales months of May and December to maximize revenue, taking advantage of seasonal buying trends.

5. Create Bundled Promotions:

- Bundle fragrance and haircare products together to drive sales, leveraging the strong correlation between these categories.

6. Implement a Recommender System:

- Develop and integrate a recommender system to provide personalized product suggestions and discounts, enhancing customer experience and driving additional sales.

THANK YOU

