Customer Segmentation & Product Recommendation

MADT 8101 - Customer Analytics

DQ
Near Home

Table of contents

DQ NearHome

01

Data Preparation

Data Preparation

Workflow overview

Data Source

Data warehouse
Google Cloud

Google BigQuery

Member
join_year < 2023

Transaction
payment_date between
'1 Jul 22' and '30 Jun 23'

Sale channel
payment_date between
'1 Jul 22' and '30 Jun 23'

Data aggregation

Product recommendation

Data preparation
- Join table
- Remove row where join_year < 1986
- Remove empty row
- Transform JSON column

Clustering

Sanity check

Variable Meaning
Context
Null

Data Type
Outlier

dataiku

02

Customer
Single View

Customer Single View

No.	Field name	Meaning
1	member_id	Define "ent" from member data as primary key
2	original_status	-
3	sponsorship	Influential power
4	registered_year	Starting year of being member
5	registered_month	Starting month of being member
6	member_duration	Membership age since starting until current

Customer Single View

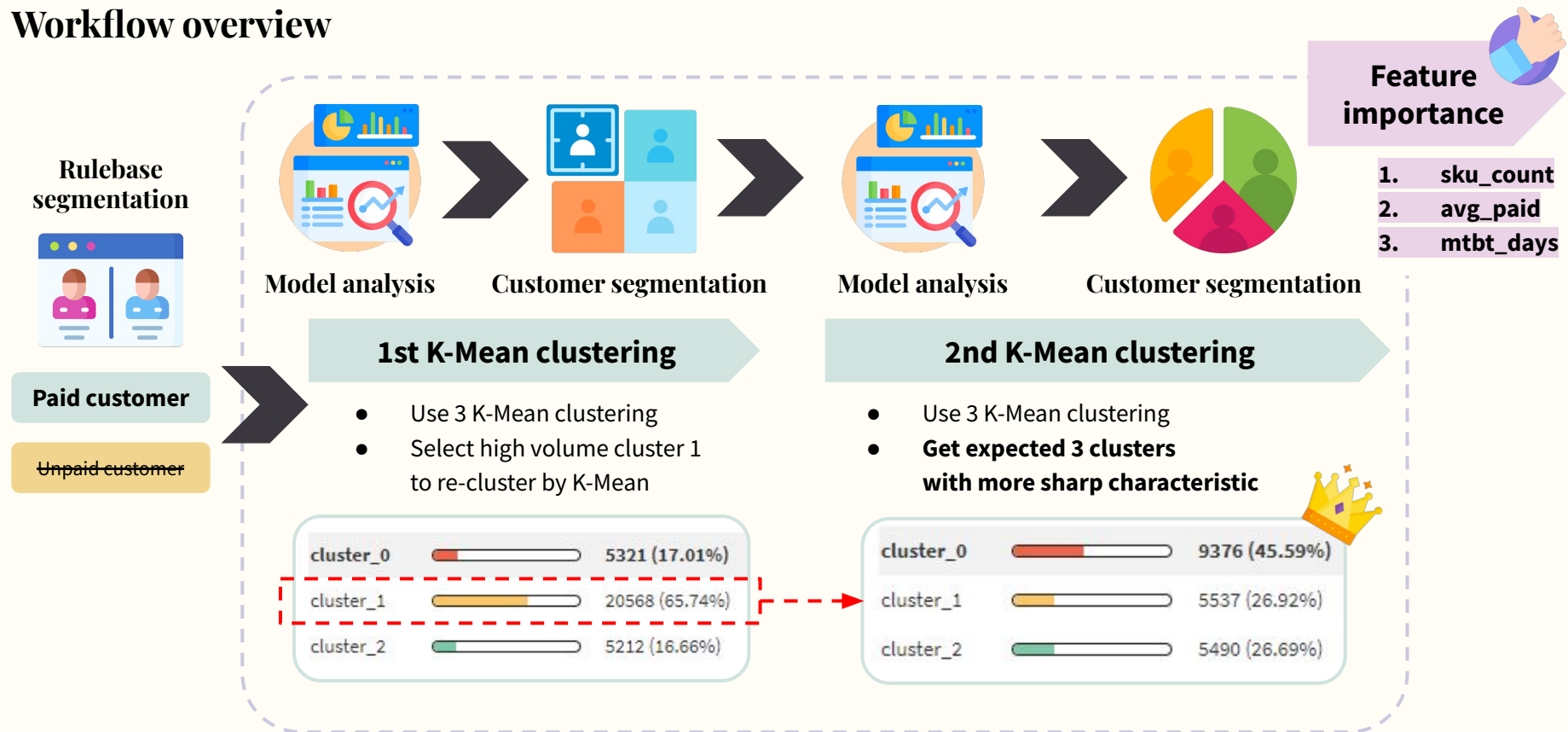| No.	Field name	Meaning
6	no_of_purchasing	Number of purchasing (Ticket counted)
7	total_spending	Sum of sales amount
8	avg_spending	Average total_spending per Ticket
9	total_product_qty	Sum product quantity
10-35	total_product_qty_bycategory	Sum product quantity by category (26 category)
36	category_count	Count purchased category
36	sku_count	Count purchased sku
38	Percent_paid	Paid/Total_spending
39	avg_paid	Percent_paid/no_of_purchasing
40	MTBP_Days	(Last purchased-First purchased)/no. of purchasing
41-66	NO_CAT_rank	Percentile of number of purchasing by category (26 category)
67	Percent_no_online	total online transactions/total transactions
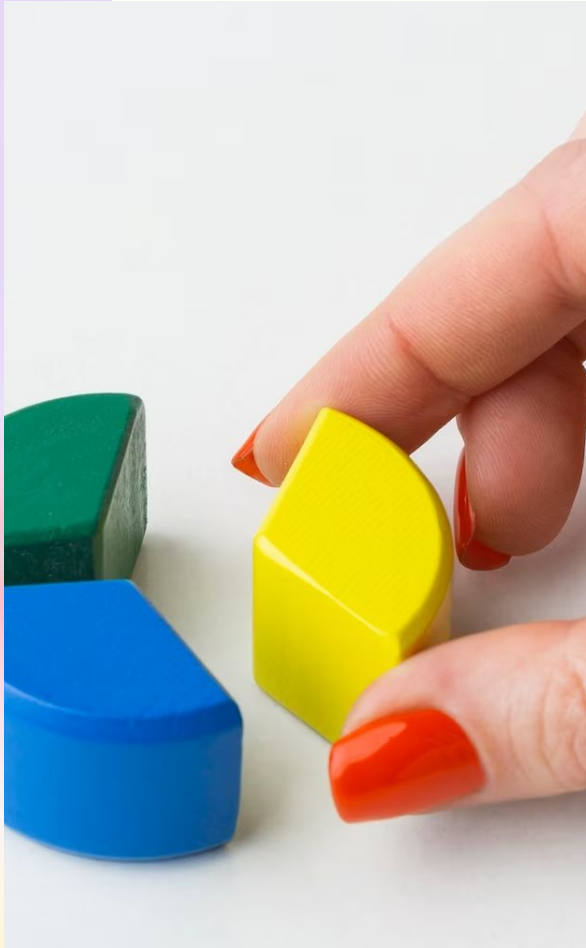
Modeling 1 – Customer segmentation

Workflow overview

Rulebase segmentation

Paid customer

~~Unpaid customer~~

Model analysis

Customer segmentation

Model analysis

Customer segmentation

1. sku_count
2. avg_paid
3. mtbt_days

1st K-Mean clustering

- Use 3 K-Mean clustering
- Select high volume cluster 1 to re-cluster by K-Mean

cluster_0		5321 (17.01%)
cluster_1		20568 (65.74%)
cluster_2		5212 (16.66%)

2nd K-Mean clustering

- Use 3 K-Mean clustering
- **Get expected 3 clusters with more sharp characteristic**

cluster_0		9376 (45.59%)
cluster_1		5537 (26.92%)
cluster_2		5490 (26.69%)

Modeling 1 – Customer segmentation

1st Cluster characteristic interpretation

Shopper Elite

Product Lover

General shopper

Modeling 1 – Customer segmentation

1st Cluster characteristic interpretation

Customer contribution **"16.66%"**

Highlight behavior

"High spending"

- Shop visit independent
- Low product variety purchased

Shopper Elite

Customer contribution **"17.01%"**

Highlight behavior

"High product variety purchased"

- Low shop visiting
- Low spending

Product Lover

Due to following reasons, we selected this cluster to re-cluster to get more sharp customer characteristic
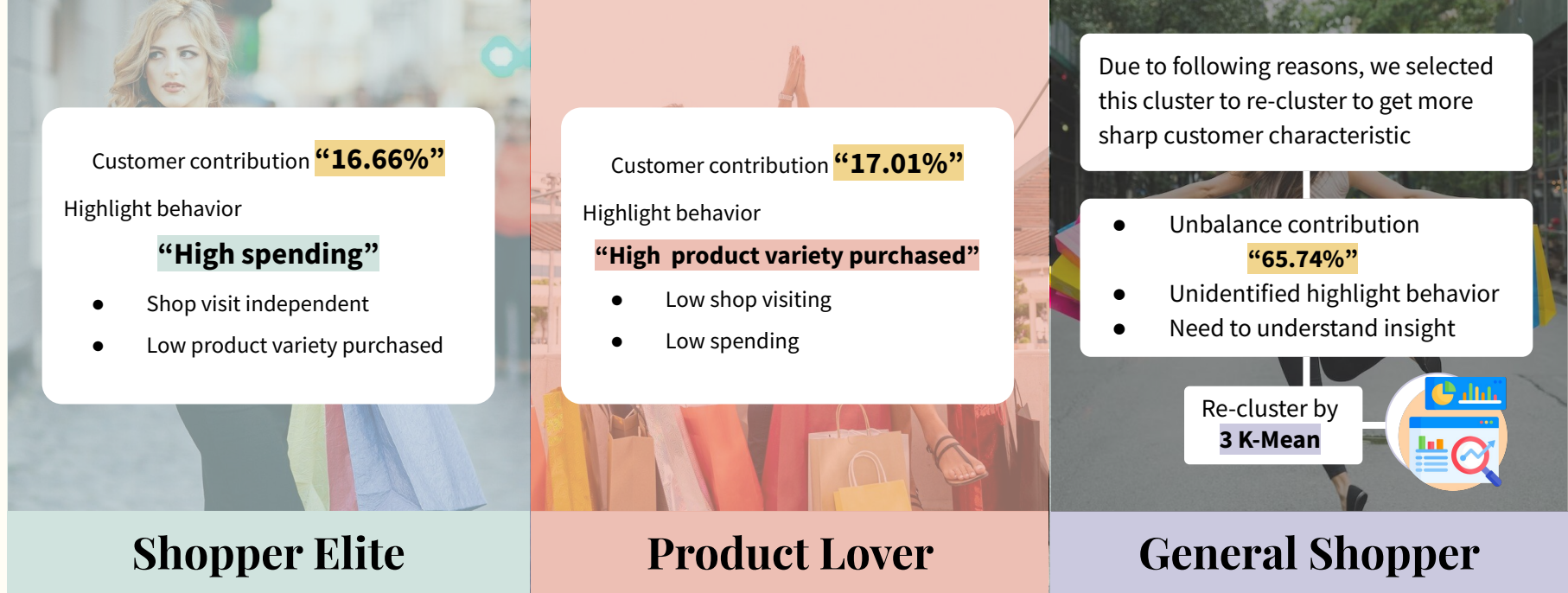
- Unbalance contribution **"65.74%"**
- Unidentified highlight behavior
- Need to understand insight

Re-cluster by **3 K-Mean**

General Shopper

Modeling 1 – Customer segmentation

2nd Cluster characteristic interpretation

Modeling 1 – Customer segmentation

2nd Cluster characteristic interpretation

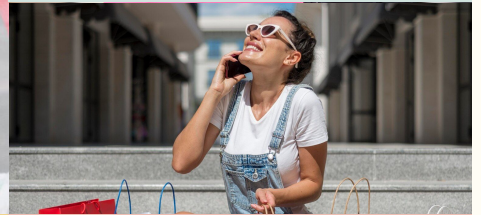Customer contribution **"26.69%"**

Highlight behavior

- Low shop visiting
- Low spending
- Low product variety purchased

Need Attentions

Customer contribution **"26.92%"**

Highlight behavior

- **"High spending"**
- **"High visiting"**
- Low product variety purchased

Shopper Loyalist

Customer contribution **"45.59%"**

Highlight behavior

- **High product variety purchased**
- Low spending
- Low shop visiting

Product Leader

Modeling 1 – Customer segmentation

Focus group criteria selection

Customer segmentation	Member	AVG Spending	AVG SKU	AVG MTBT/Days
Product Leader	9.4K	153K	10.51	21.11
Shopper Loyalist	5.5K	514.1K	7.04	26.31
Need Attentions	5.5K	231.5K	4.51	64.09
Product Lover	5.3K	195.4K	28.15	14.61
Shopper Elite	5.2K	1.2M	3.66	126.36

Objective: Increase sales revenue

Strategy: Up-selling promotion

- Focus on high spending group
- Approach up-selling promotion
- Marketing personalization through
  - product recommendation

04

Product Recommendation

Product Recommend

Products by Categories
Top 10 Purchased

Customer contribution **"16.66%"**

Highlight behavior

"High spending"

- Shop visit independent
- Low product variety purchased

Shopper Elite

CAT.Products		Quantity
1.	qty_8	5,824
2.	qty_K	4,892
3.	qty_5	4,789
4.	qty_6	2,767
5.	qty_B	2,554
6.	qty_0	1,287
7.	qty_X	1,087
8.	qty_7	1,049
9.	qty_4	529
10.	qty_Y	448

Modeling – Product Recommend by Machine Learning

K-mean segmentation

Shopper Elite

Shopper Loyalist

Entity of Users
Categories of Product
- CAT_A
- CAT_B
- CAT_C
- And More...

Collaborative Filtering (Content-Based)

Recommend

Content of
A Similar to B
Recommend "B" if "A" had been purchased

Upselling

High Spending
High Visiting
More Variety Products
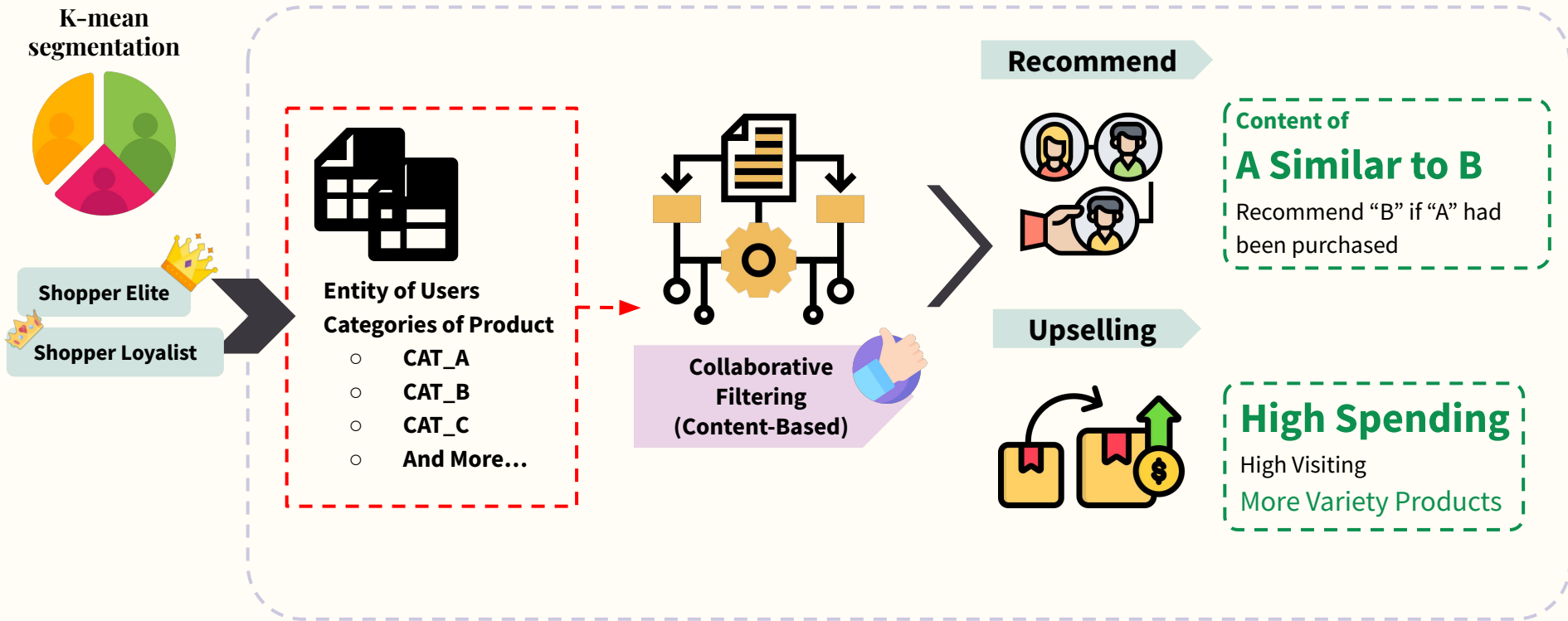
Modeling – Product Recommend by Machine Learning

Shopper Elite

	qty_B	qty_6	qty_8	qty_K	qty_0	qty_2	qty_X
qty_B	1	0.240118	0.342686	0.1476	0.147991	0.080042	0.136784
qty_6	0.240118	1	0.322872	0.311202	0.316546	0.266084	0.221972
qty_8	0.342686	0.322872	1	0.268932	0.265927	0.143441	0.261665
qty_K	0.1476	0.311202	0.268932	1	0.240581	0.123563	0.254526
qty_0	0.147991	0.316546	0.265927	0.240581	1	0.135623	0.19536
qty_2	0.080042	0.266084	0.143441	0.123563	0.135623	1	0.082241
qty_X	0.136784	0.221972	0.261665	0.254526	0.19536	0.082241	1
qty_7	0.544027	0.286822	0.163813	0.143991	0.083151	0.091272	0.115894
qty_E	0.026502	0.067784	0.057588	0.33672	0.06069	0.026374	0.134616
qty_U	0.015445	0.029533	0.02076	0.017833	0.015483	0.020684	0.028917
qty_5	0.144739	0.394246	0.205038	0.204347	0.174903	0.201869	0.144766
qty_L	0.006085	0.026178	0.023245	0.064666	0.022874	0.0	0.018309
qty_4	0.166493	0.078446	0.2629	0.119487	0.113664	0.055474	0.127842
qty_S	0.070058	0.349218	0.208525	0.135623	0.166061	0.147661	0.102183
qty_D	0.175927	0.158501	0.161428	0.135642	0.125267	0.037418	0.085321

Shopper Loyalist

	qty_	qty_6	qty_8	qty_K	qty_0	qty_2	qty_X
qty_B	0. 198	0.323569	0.385534	0.364681	0.247454	0.279945	
qty_6	0.519	1	0.446702	0.449569	0.43098	0.394212	0.349612
qty_8	0.323569	0.446702	1	0.298496	0.247574	0.178831	0.265223
qty_K	0.385534	0.449569	0.298496	1	0.405149	0.238563	0.389839
qty_0	0.364681	0.43098	0.247574	0.405149	1	0.222933	0.277032
qty_2	0.247454	0.394212	0.178831	0.238563	0.222933	1	0.182314
qty_X	0.279945	0.349612	0.265223	0.389839	0.277032	0.182314	1
qty_7	0.423804	0.40 712	0.174864	0.232311	0.228283	0.1753	0.164062
qty_E	0.065659	0.05693	0.041586	0.269057	0.083817	0.013242	0.125163
qty_U	0.141129	0.120655	0.091197	0.064045	0.088048	0.047289	0.046698
qty_5	0. 34	0.535287	0.243337	0.389387	0.26146	0.344121	0.343234
qty_L	0.091346	0.100254	0.059369	0.109617	0.121131	0.056124	0.054218
qty_4	0.204418	0.281927	0.187326	0.195121	0.177498	0.130907	0.162549
qty_S	0.256293	0.331143	0.175574	0.311144	0.290348	0.210722	0.205578
qty_D	0.259857	0.29398	0.18546	0.207485	0.164215	0.127513	0.173916

05

Business
Action

Business action

Customer

Purchased

Is existing group ?

*Upcoming segmentation roadmap

Customer segmentation

Shopper Elite

Shopper Loyalist

Do purchased product match to recommendation list ?

Product recommendation

Recommend product

Utilize history sales data backward 1 Month

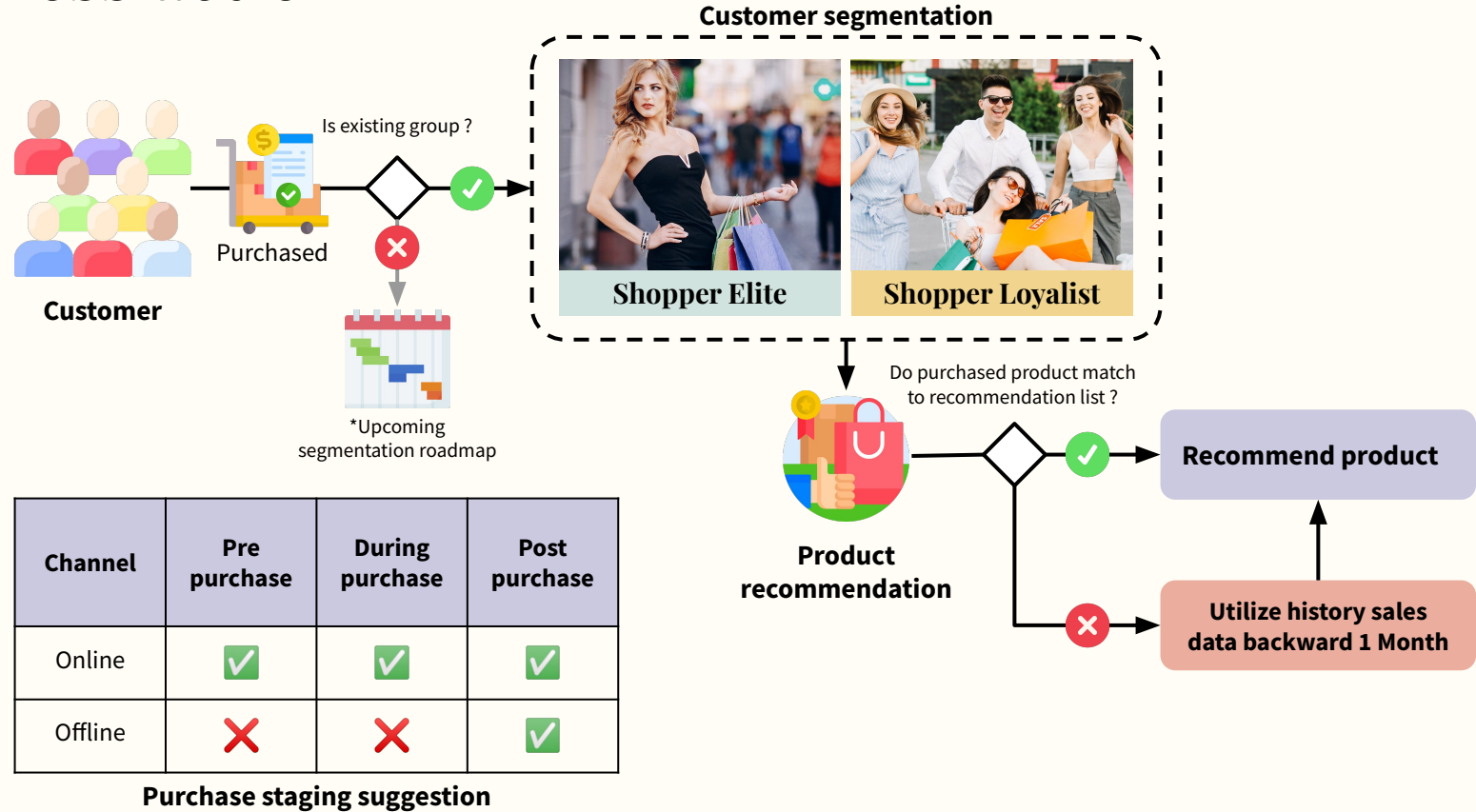| Channel	Pre purchase	During purchase	Post purchase
Online	✅	✅	✅
Offline	❌	❌	✅

Purchase staging suggestion

Thank You 🙂

Project Team Member

- Kanin Jaichean 6510424015
- Alongkon Sewaleekunanon 6510424017
- Anuwat Kumkaew 6510424022
- Apiwat Iamborisuth 6510424026