



HDI HOLDINGS

SEGMENTATION

PRODUCT RECOMMENDATION

MADT8101 CUSTOMER ANALYTICS





1	2	3	4	5	6
					
Business objective	Dataset and Limitation	Preparation and feature extraction	Modeling Segmentation	Modeling Product recommendation	Recommendation

1

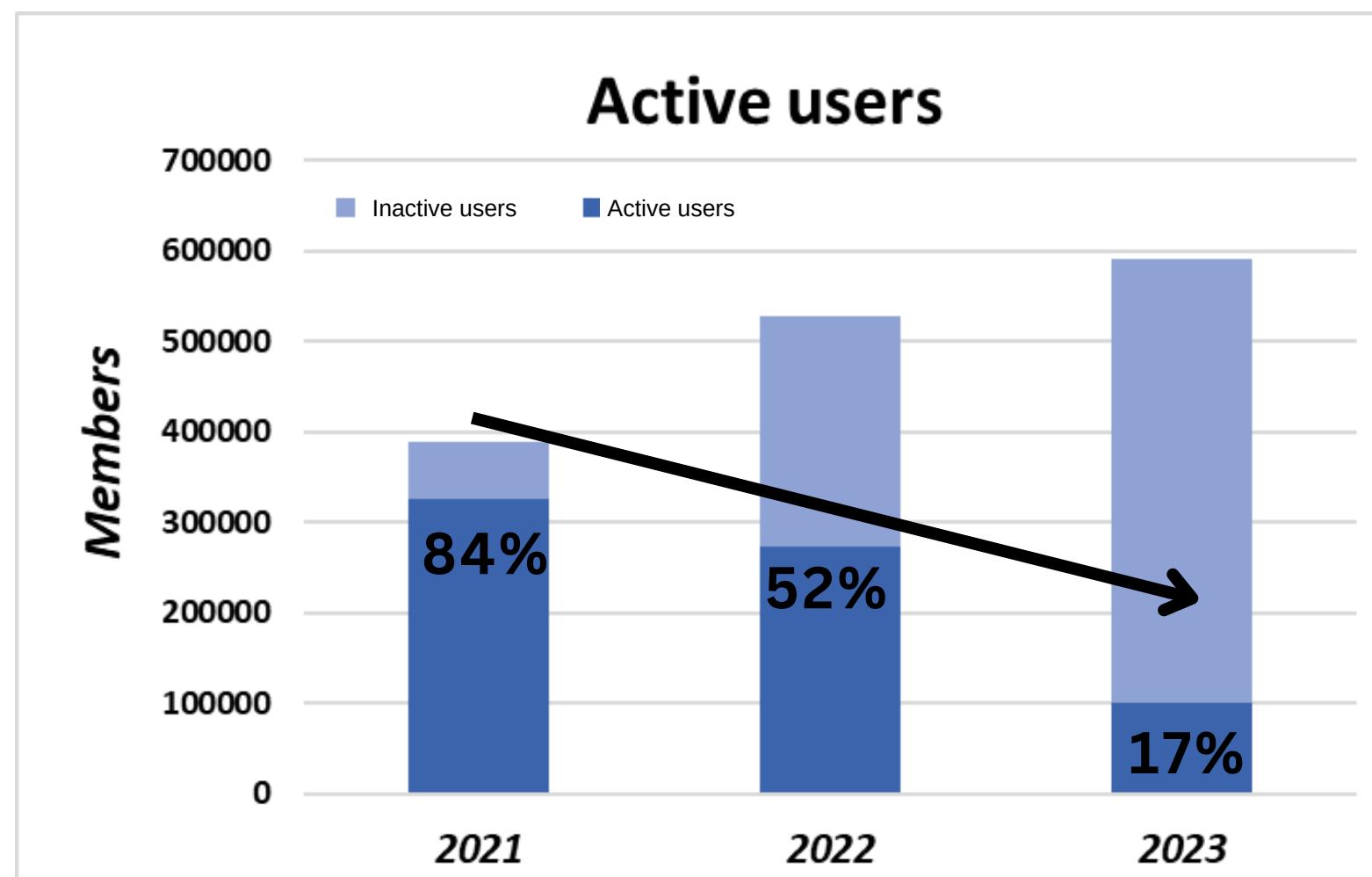
BUSINESS OBJECTIVE



BUSINESS OBJECTIVE

From historical information, the business have consistent growth in the number of new member each year.

Simultaneously, they are also have decline in active customer. Therefore, the company may initiate segmentation member to analysis and plan strategy for further action plan and also perform product recommendation for increase opportunities of transaction growth.



Segmentation



Segmentation enables to deliver personalized and targeted messages to our customers following by level. This personalized approach builds stronger connections, increases engagement, and ultimately boosts sales

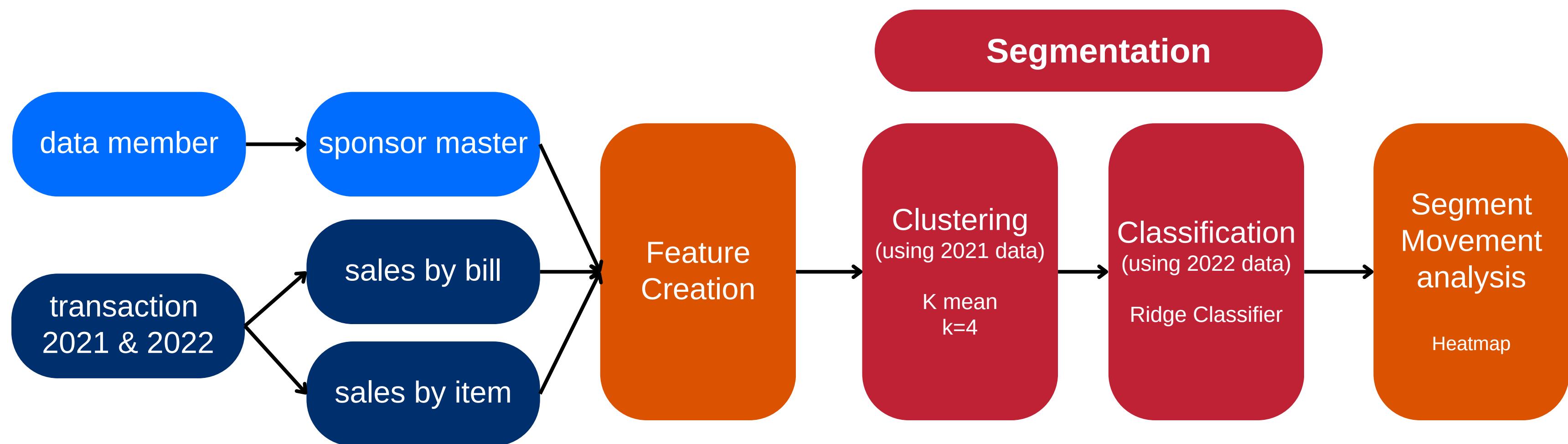
Product recommendation



To analyze customers' patterns and preferences and provide suitable market strategies to increase sale, improve customer experience and satisfaction, such as Upselling and Cross-selling

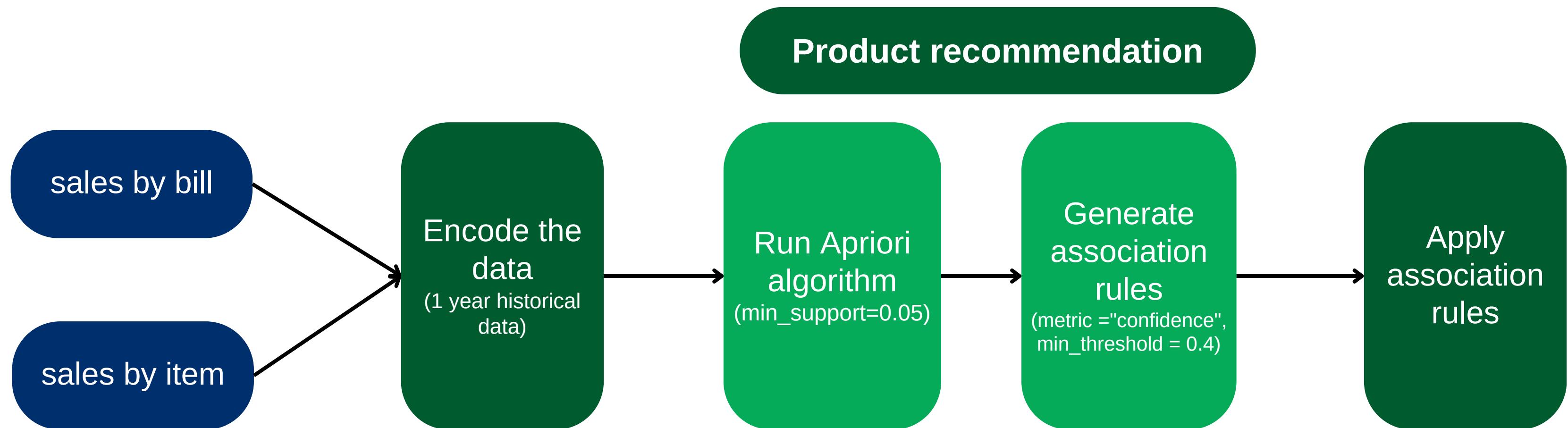


SEGMENTATION - PROJECT OVERVIEW





PRODUCT RECOMMENDATION - PROJECT OVERVIEW

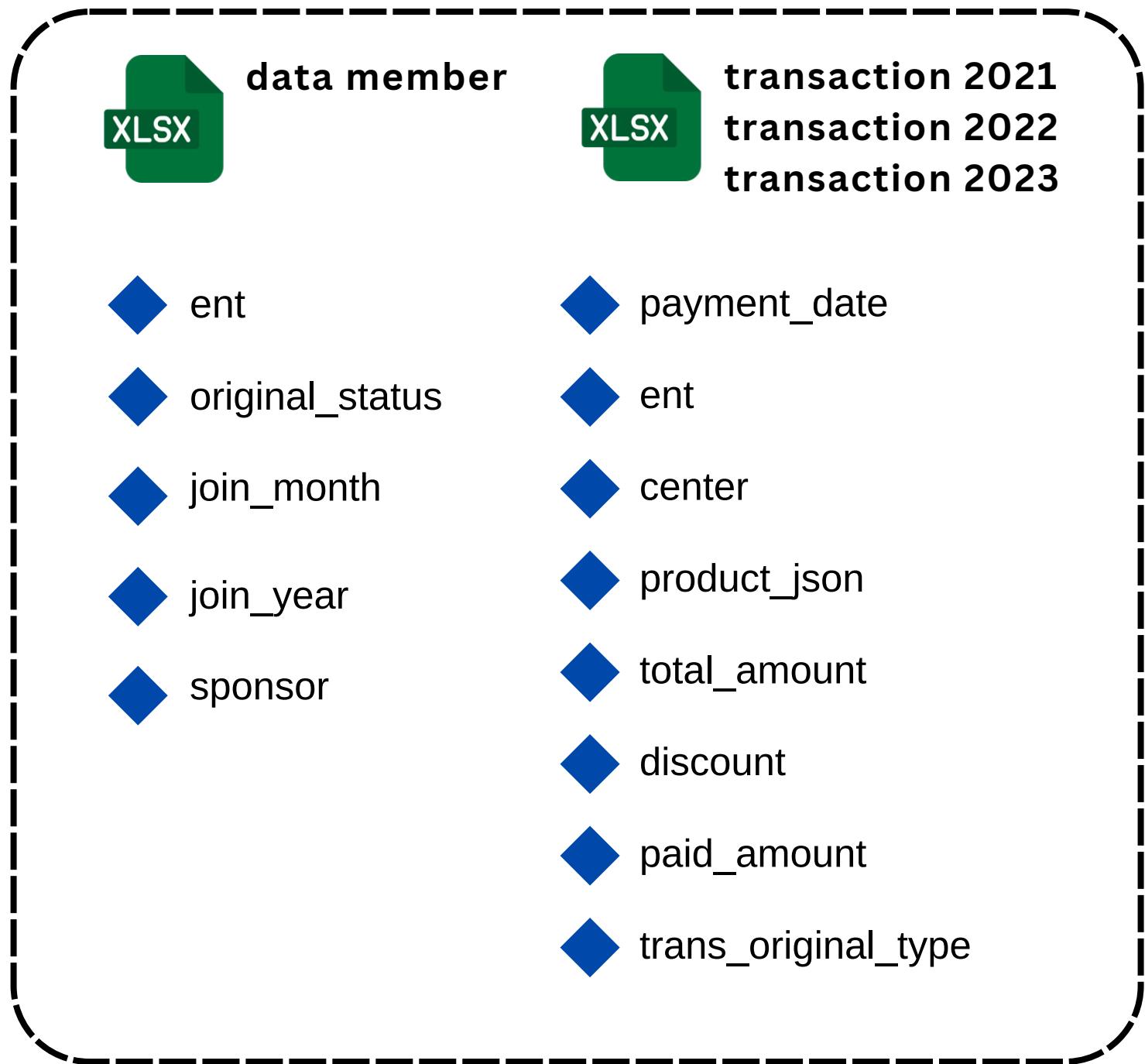


2

DATASET AND LIMITATION



DATASET



LIMITATION

-
- Cannot identify meaning of column
 - original_status
 - paid_amount
 - center
 - Cannot verify correctness of information
 - Outliner joined year (before business' established year :1986)
 - "ent" joined network before "sponser"
 - Fold data
 - Cannot identify item price per unit
 - Cannot identify pattern discount

3

PREPARATION AND FEATURE EXTRACTION





PREPARATION AND FEATURE EXTRACTION

MEMBER



- Replace join years before 1986 with 1986 and create lifetime of member
 - Lifetime
- Extract one-hot encoding on "original_status"
 - Status A - O
- Grouping by "sponsor" to calculate number of "ent"
 - No. of downline

ITEM



- Unfold and Unnest item list
- Create item profile (purchasing frequency and unique item) for each "ent"
 - Unique item (Item variety)
 - Total Qty (Volume)
 - % of purchase frequency (Repurchase portion)
 - Unique/Frequency by Channel

TRANSACTION



- Grouping by amount to calculate performance of each "ent"
 - Total / Avg. Amount spending
 - Value/Item by channel
- Calculate frequency and recency of each "ent"
 - Last purchase
 - Purchase frequency
 - No. of ticket by channel

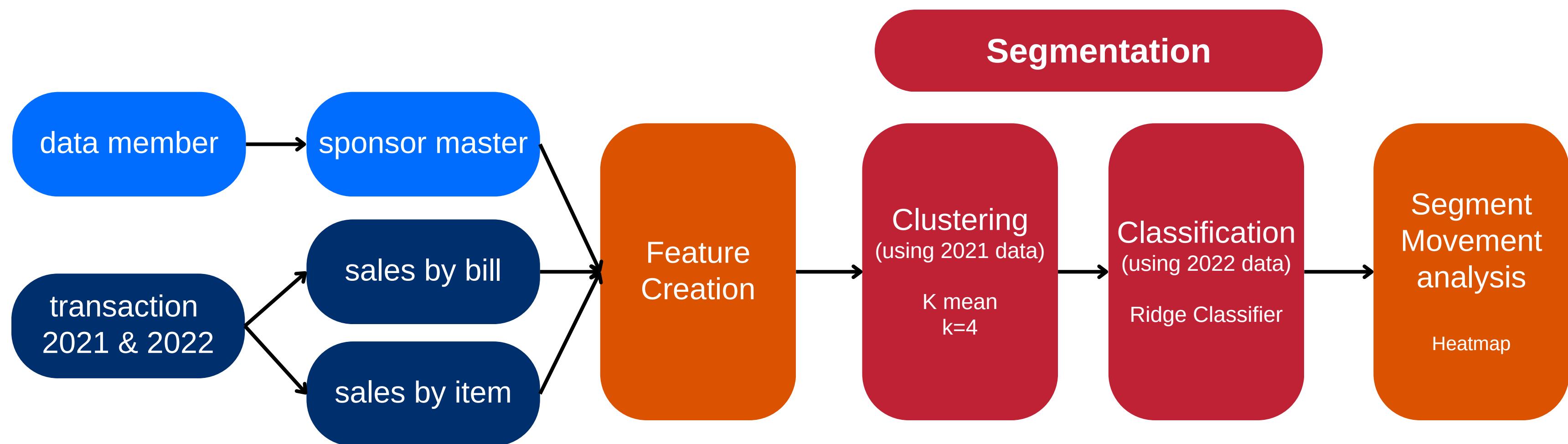
4

MODELING

- SEGMENTATION

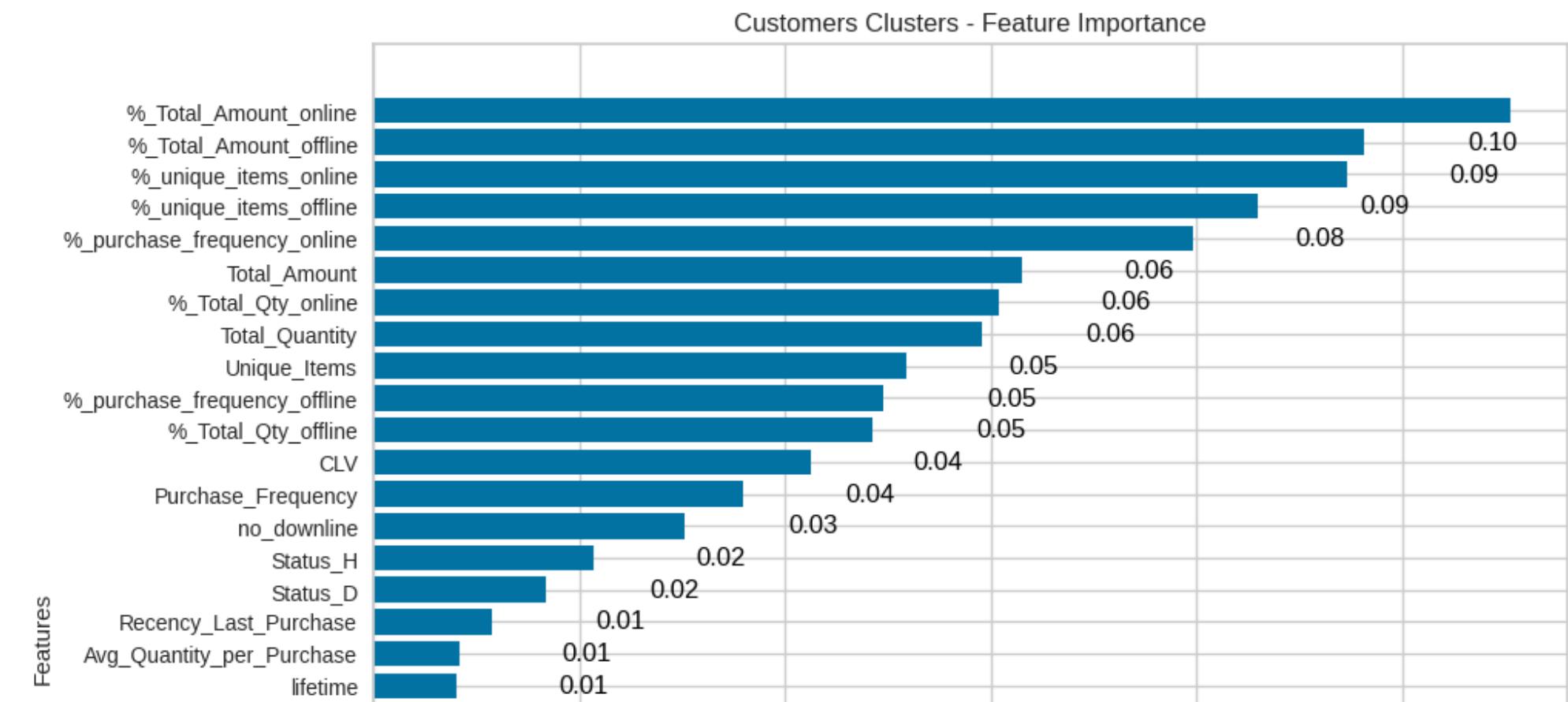
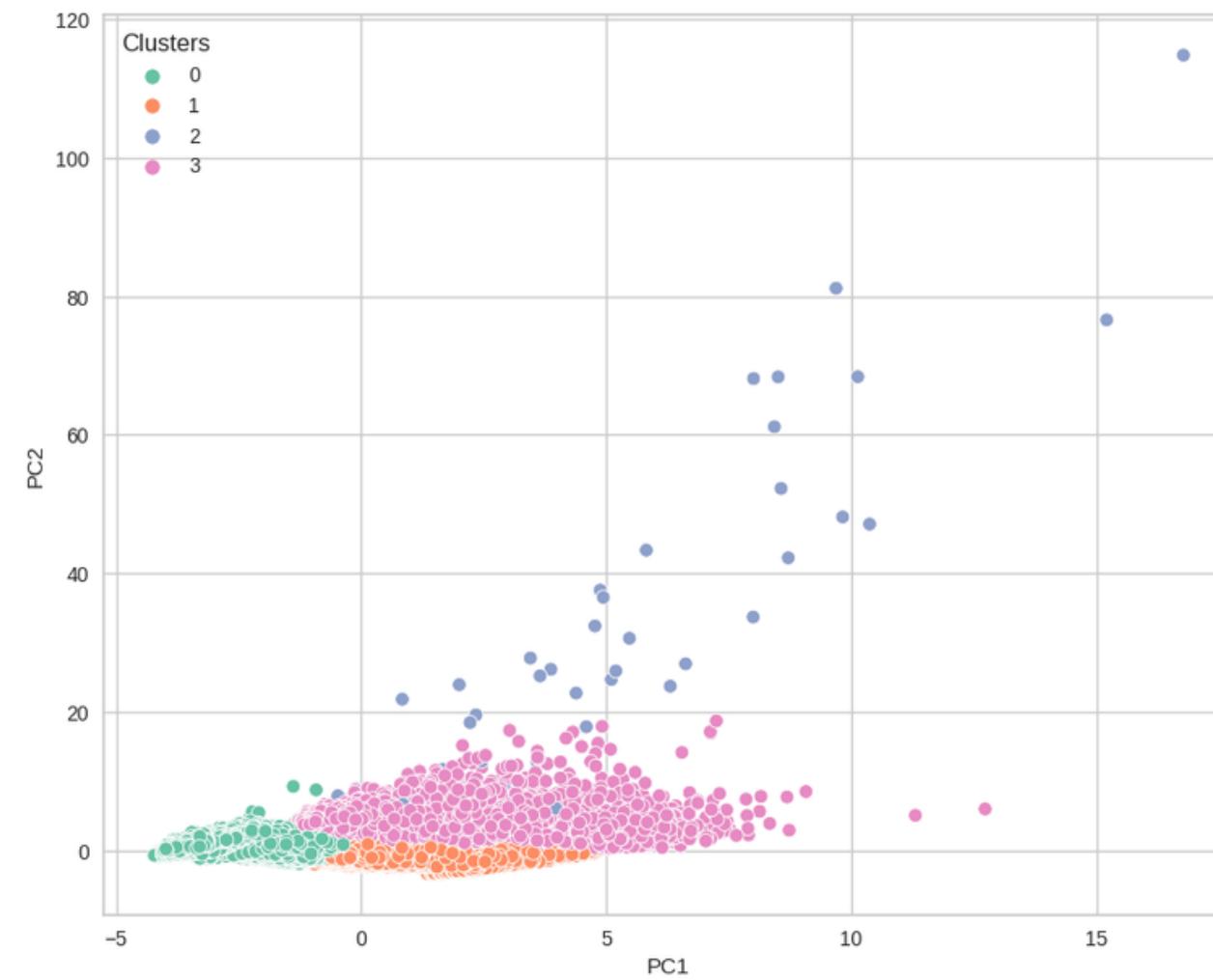


SEGMENTATION - PROJECT OVERVIEW



SEGMENTATION - CLUSTERING >> Y21

Before										
Clusters	Cluster Name	Total_Amount	%_Total_Amount_online	%_Total_Amount_offline	Unique_Items	Purchase_Frequency	Total_Qty	CLV	no_downline	
0	Star (Offline)	3,456,180.56	8%	91%	4.44	3.33	8.32	5,589,269.35	4.00	
1	Star (Online)	4,177,173.94	87%	13%	5.67	4.16	9.53	3,756,340.28	3.77	
2	Diamond	21,832,055.71	30%	70%	17.84	28.26	2,687.46	273,499,251.43	52.00	
3	Gold	13,836,763.92	49%	51%	14.67	13.49	38.41	45,951,916.10	18.41	





SEGMENTATION - CLASSIFICATION >> Y22

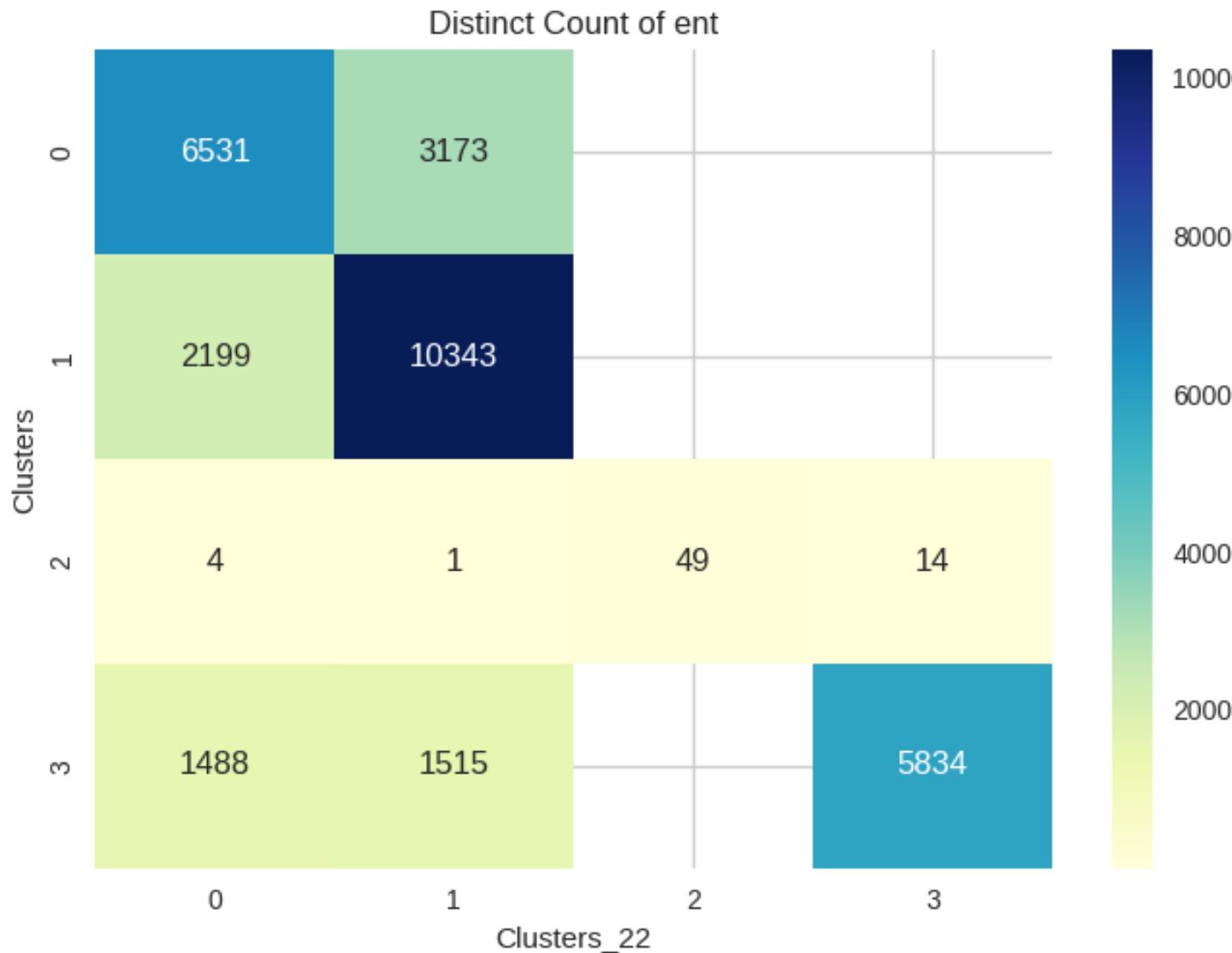
After

Clusters_22	Cluster Name	Total_Amount	%_Total_Amount_online	%_Total_Amount_offline	Unique_Items	Purchase_Frequency	Total_Qty	CLV	no_downline
0	Star (Offline)	6,811,551.16	14%	86%	8.75	6.35	20.20	20,457,862.76	6.40
1	Star (Online)	5,442,216.92	92%	8%	8.10	5.88	17.20	10,628,265.62	4.24
2	Diamond	34,748,157.88	40%	60%	33.96	44.88	1,569.21	377,871,769.23	53.85
3	Gold	19,884,353.66	61%	39%	22.00	20.33	70.46	97,991,703.73	19.30





SEGMENT MOVEMENT ANALYSIS



- No. customer Y21 : 48K
- No. customer Y22 : 57K

Before										
Clusters	Cluster Name	Total_Amount	%_Total_Amount_online	%_Total_Amount_offline	Unique_Items	Purchase_Frequency	Total_Qty	CLV	no_downline	
0	Star (Offline)	3,456,180.56	8%	91%	4.44	3.33	8.32	5,589,269.35	4.00	
1	Star (Online)	4,177,173.94	87%	13%	5.67	4.16	9.53	3,756,340.28	3.77	
2	Diamond	21,832,055.71	30%	70%	17.84	28.26	2,687.46	273,499,251.43	52.00	
3	Gold	13,836,763.92	49%	51%	14.67	13.49	38.41	45,951,916.10	18.41	

After										
Clusters_22	Cluster Name	Total_Amount	%_Total_Amount_online	%_Total_Amount_offline	Unique_Items	Purchase_Frequency	Total_Qty	CLV	no_downline	
0	Star (Offline)	6,811,551.16	14%	86%	8.75	6.35	20.20	20,457,862.76	6.40	
1	Star (Online)	5,442,216.92	92%	8%	8.10	5.88	17.20	10,628,265.62	4.24	
2	Diamond	34,748,157.88	40%	60%	33.96	44.88	1,569.21	377,871,769.23	53.85	
3	Gold	19,884,353.66	61%	39%	22.00	20.33	70.46	97,991,703.73	19.30	

Movement Analysis:

- Successfully increase purchased quantity through increase in unique items in both Star group (online and offline)
- Drop-off of Gold (3) by 36%, transitioning to become stars rank (0, 1).
- Those who remained as gold level have doubled their purchases.
- Contrastingly, Diamond's spending has declined by 42%.
- The decrease in Gold level resulted in a 26% reduction in the number of Diamond, from 70 to 52 in 2022.

5

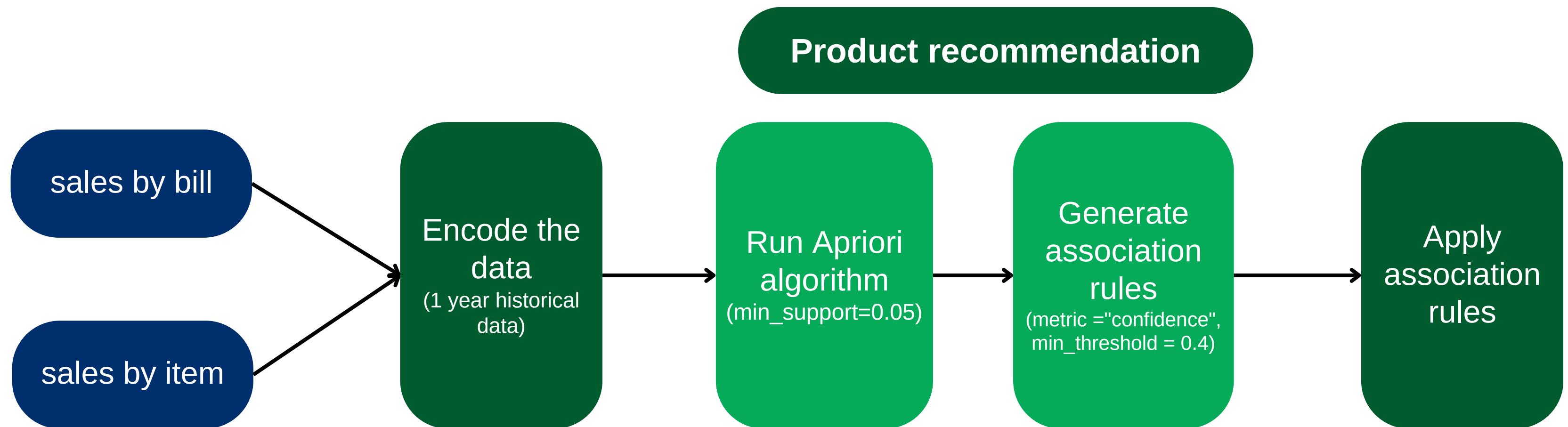
MODELING



- PRODUCT
RECOMMENDATION



PRODUCT RECOMMENDATION - PROJECT OVERVIEW





PRODUCT RECOMMENDATION - TOP 5 ITEMS IN SUPPORT VALUE



All Customers

Item	Support Values
(5C4CCE)	0.317857
(5C4CC4)	0.236226
(6CQC41)	0.225366
(8C4CCR)	0.177295
(XC4CCW)	0.166011

Online Customers

Item	Support Values
(6CQC41)	0.449441
(5C4CCE)	0.200716
(5C4CC4)	0.161674
(5C4C44)	0.155814
(7C4CC4)	0.152903

Offline Customers

Item	Support Values
(6CQC41)	0.301466
(5C4CCE)	0.230137
(5C4C44)	0.187512
(5C4CC4)	0.186000
(7C4CC4)	0.163892



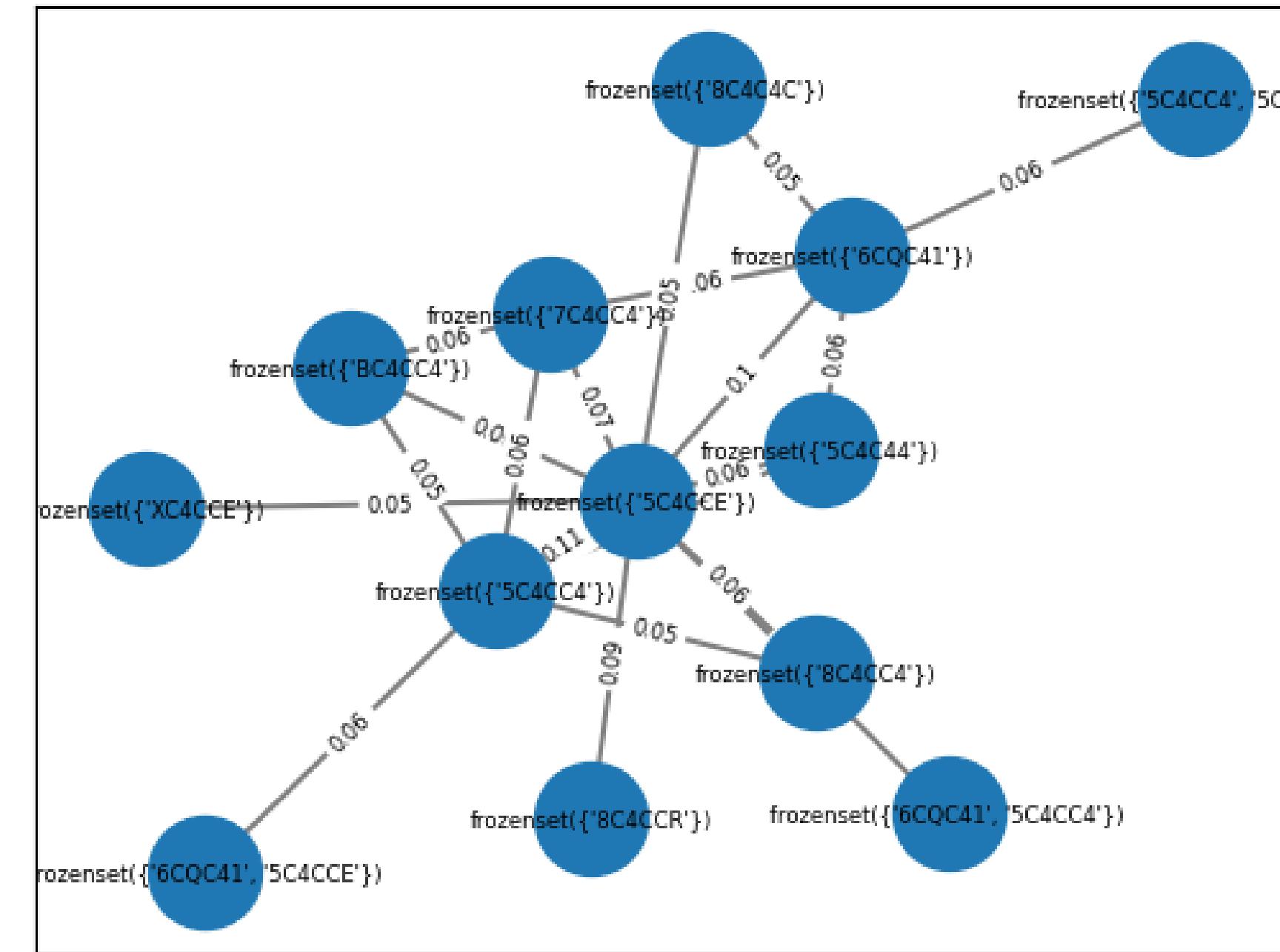
PRODUCT RECOMMENDATION - MARKET BASKET RECOMMENDATION

All Customers : Top 5 in total support

antecedents	consequents	support	confidence 	lift 	leverage 
(5C4CC4)	(5C4CCE)	0.110338	0.467086	1.469484	0.035252
(6CQC41)	(5C4CCE)	0.102799	0.456141	1.435052	0.031165
(8C4CCR)	(5C4CCE)	0.091563	0.516445	1.624773	0.035209
(7C4CC4)	(5C4CCE)	0.066922	0.475178	1.494941	0.022156
(8C4CC4)	(5C4CCE)	0.06116	0.484526	1.524351	0.021038

PRODUCT RECOMMENDATION - MARKET BASKET RECOMMENDATION

All Customers: Support > 0.05





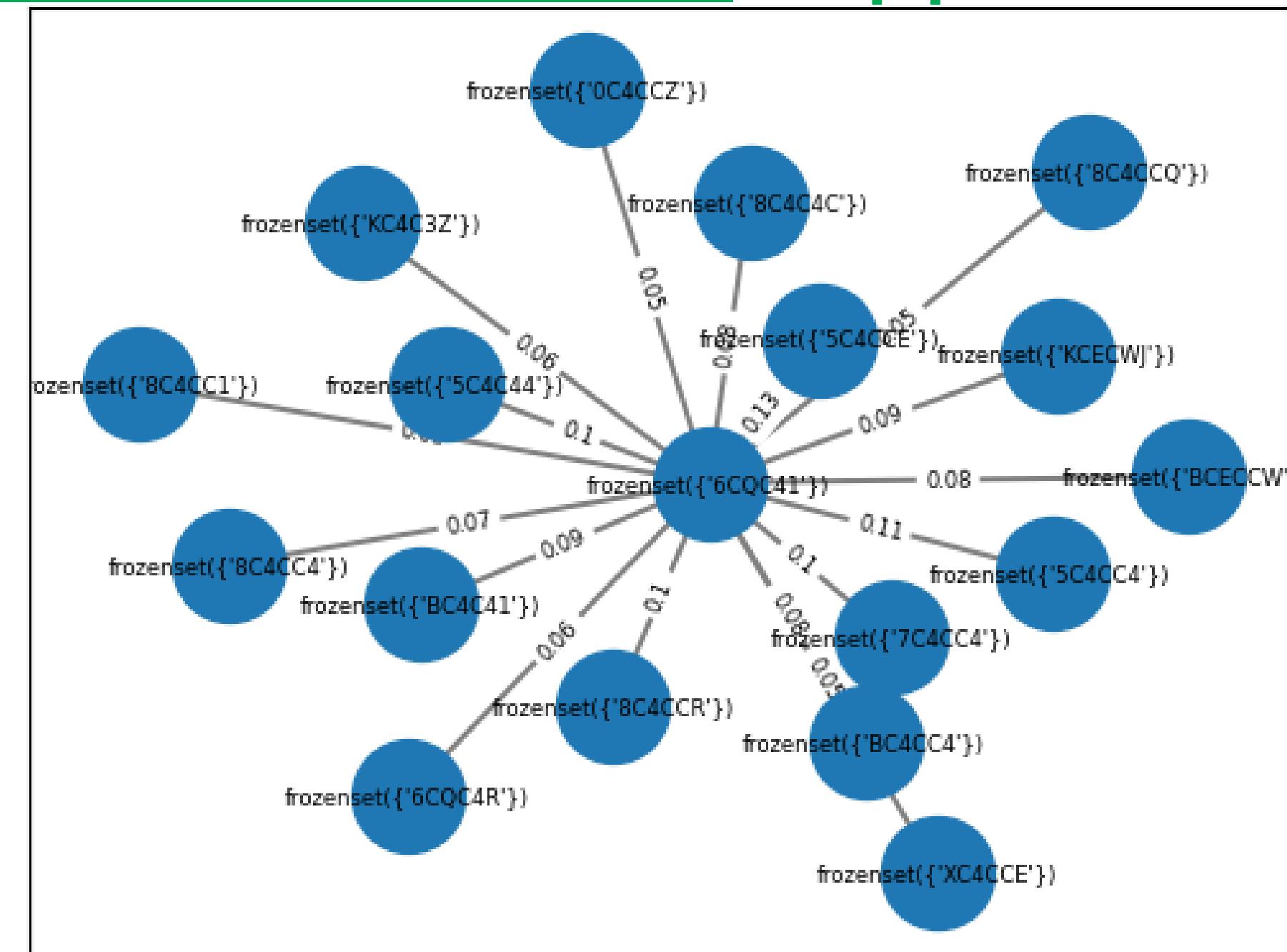
PRODUCT RECOMMENDATION - MARKET BASKET RECOMMENDATION

Online Customers: Top 5 in total support

antecedents	consequents	support 	confidence 	lift 
(5C4CCE)	(6CQC41)	0.129606	0.645716	1.436709
(5C4CC4)	(6CQC41)	0.109272	0.675879	1.503822
(5C4C44)	(6CQC41)	0.100524	0.64515	1.43545
(8C4CCR)	(6CQC41)	0.100305	0.670286	1.491377
(7C4CC4)	(6CQC41)	0.097981	0.640805	1.425782

PRODUCT RECOMMENDATION - MARKET BASKET RECOMMENDATION

Online Customers: Support > 0.05





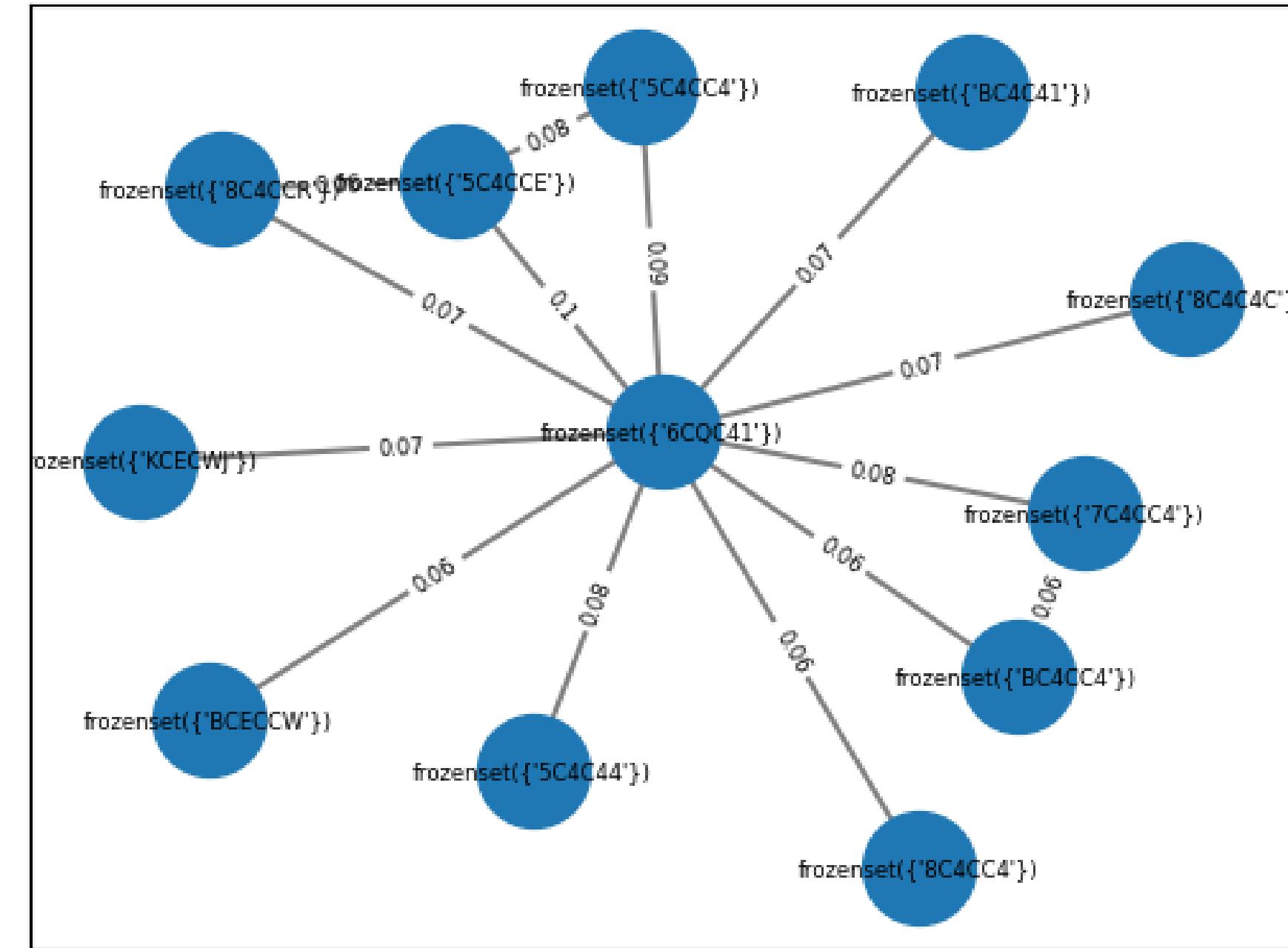
PRODUCT RECOMMENDATION - MARKET BASKET RECOMMENDATION

Offline Customers: Top 5 in total support

antecedents	consequents	support 	confidence 	lift 
(5C4CCE)	(6CQC41)	0.102387	0.444898	1.475783
(5C4CC4)	(6CQC41)	0.087801	0.47205	1.565847
(5C4C44)	(6CQC41)	0.082344	0.439138	1.456676
(7C4CC4)	(6CQC41)	0.078846	0.481089	1.595832
(5C4CC4)	(5C4CCE)	0.075175	0.404167	1.756205

PRODUCT RECOMMENDATION - MARKET BASKET RECOMMENDATION

Offline Customers: Support > 0.05



6

RECOMMENDATION



BUSINESS INSIGHT RECAP

Data Project

SEGMENTATION

PRODUCT RECOMMENDATION

DO GOOD

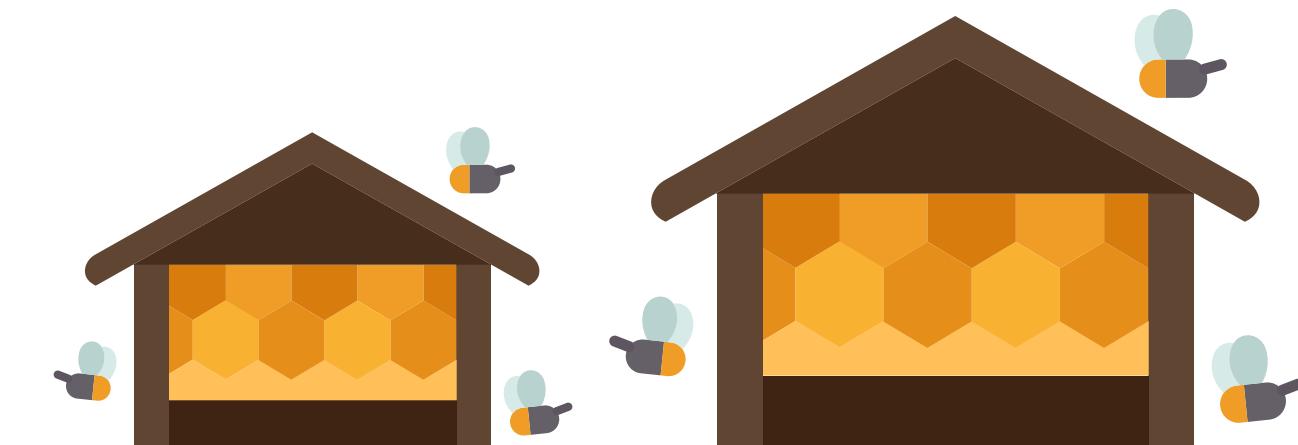
- Cross category promotion to increase unique items and sales qty
- Product Portfolio curation



DO BETTER

- Membership retention (especially in gold and diamond tier)
- Membership progress through tiers

- Explore product recommendation by customer tier and by product category





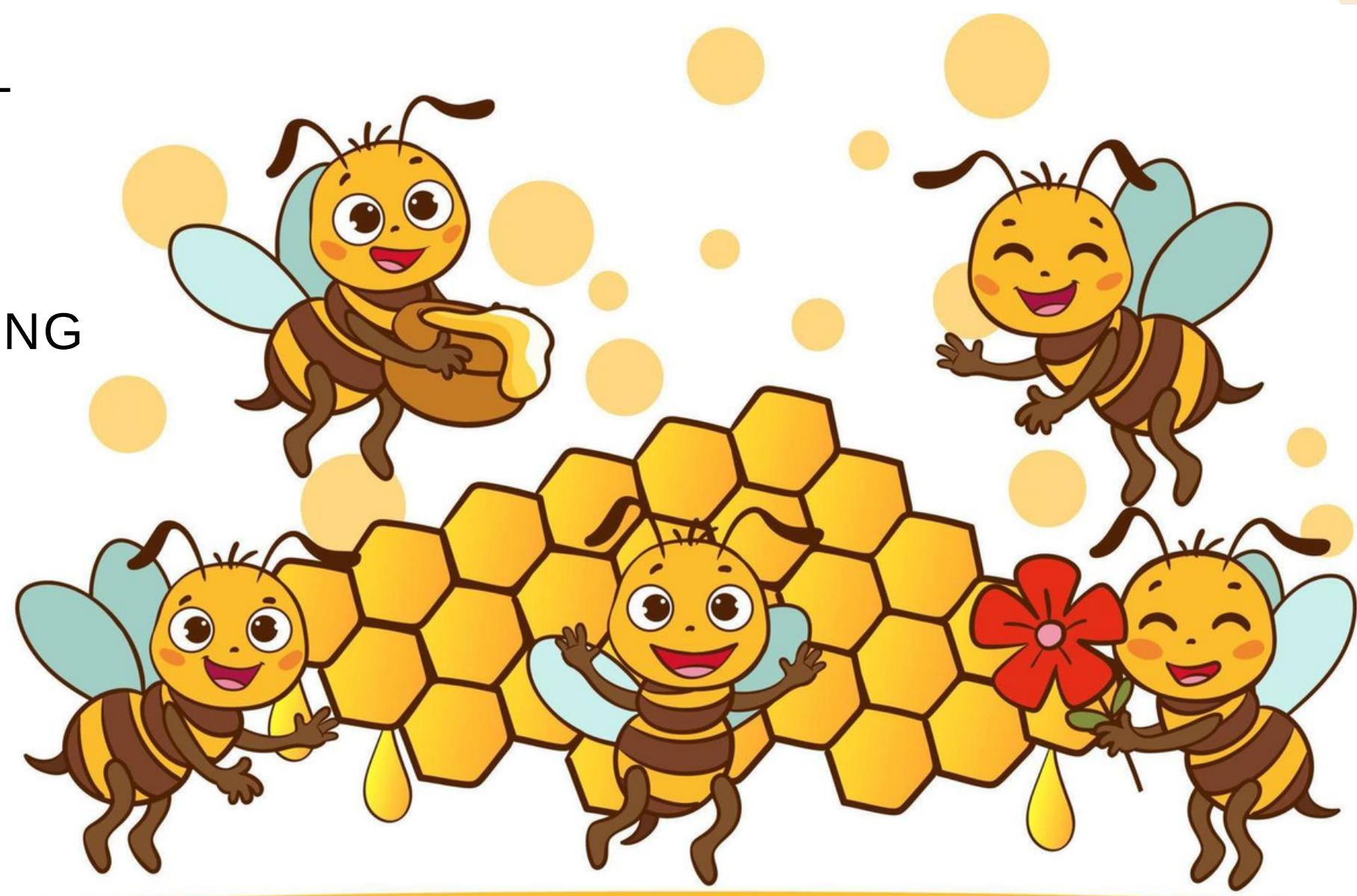
NEXT STEPS:

Topics	Business Objective	Recommended Next Steps	Responsible
Data cleansing	Create complete and accurate data as a hygiene factor	Data Cleansing for accurate prediction	DBA
Sale growth	Growing Sales through increase in spending	Explore cross category product recommendation per each segment including explore other recommendations algorithm	Data Analytics
Sale growth	Increase spending per customer	Trial product recommendations insights in both on and offline sale channels	Sales & marketing
Retained customer	Maintain existing customers	Review the membership loyalty program to start retaining Gold and Diamond member	Marketing
Customer growth	Growing customer lifetime & no. of downline	Review timeline & incentive to grow customer through existing tier scheme	Marketing

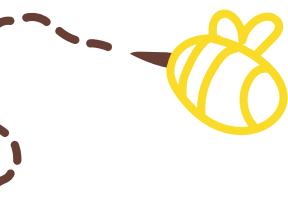


TEAM MEMBER

- ◆ 6510424006 - ANGWARA PRADASOOK
- ◆ 6510424009 - CHUTIMA KHUNVAPANICHKUL
- ◆ 6510424012 - THANIPAN SOMRANG
- ◆ 6510424018 - JANEJIRA RUSSMISAENGTHONG
- ◆ 6510424030 - CHAWIN TAWEEROJKULSRI



APPENDIX



BUSINESS INTRODUCTION





The company offers natural products from bee, they have five main categories, namely supplements for adult & Kid, personal care, skincare, food and beverages.

HDI was started in 1986 by Mr. Peter Chia, who desired to provide a better life for his family. HDI emerged as a trailblazer in the development of the **Network Marketing Business** in Asia, specifically in Singapore, Malaysia, Indonesia, Hong Kong, and the Philippines.



CATEGORY & PRODUCT



HDI BEE PROPOLIS



HDI KIDS KIDS3



BEE BOTANICS™ FACIAL WASH GEL



HDI NATURALS PROLIZ



BSKIN VITA ADVANCED LINE

HEALTH SUPPLEMENTS

- Vitamin

HEALTH SUPPLEMENTS FOR KID

- Vitamin

NATURAL PERSONAL CARE

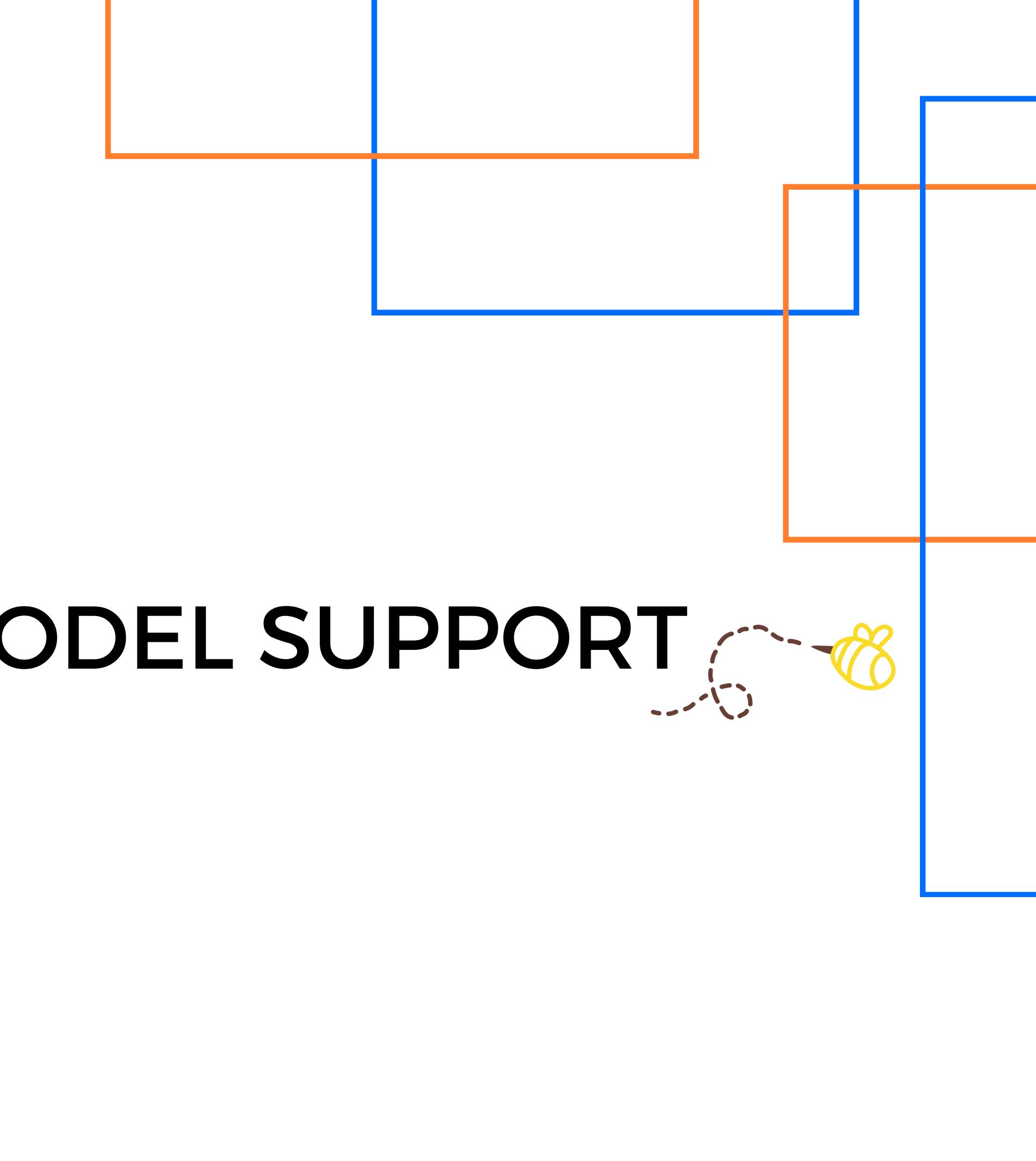
- Toothpaste
- Facial wash
- Body lotion
- Shampoo

HEALTHY FOOD AND DRINK

- Candy

SKINCARE

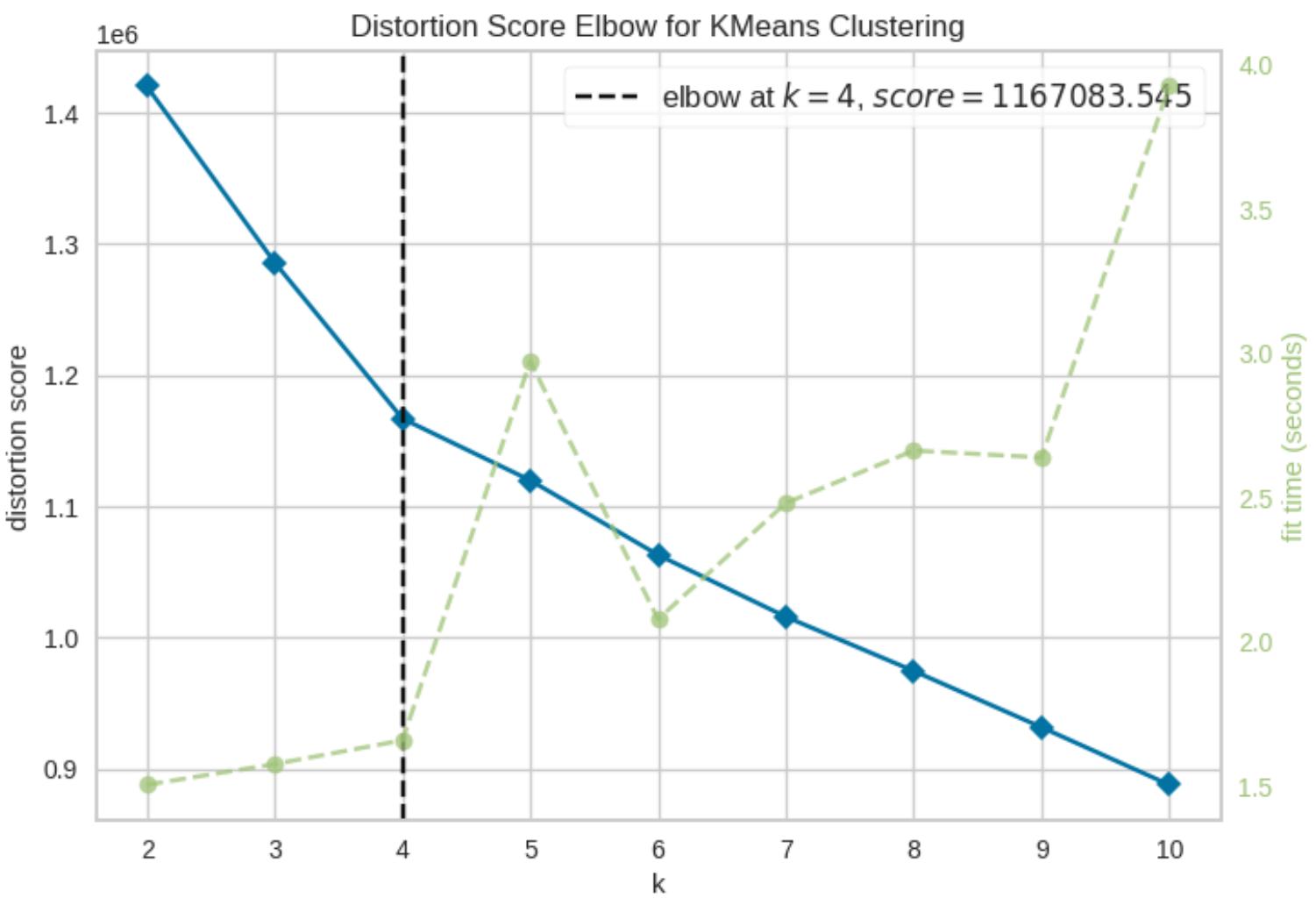
- Powder
- Sunscreen
- Serum
- Treatment
- Cleanser



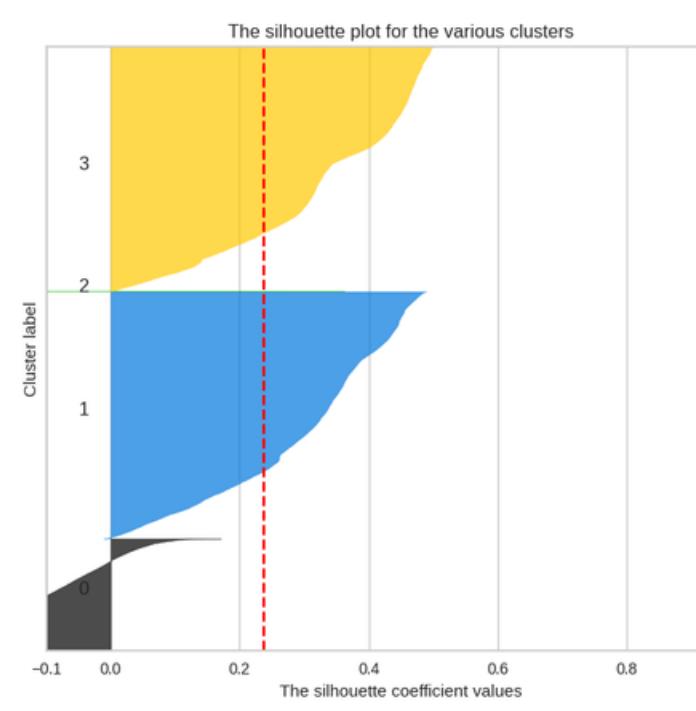
MODEL SUPPORT



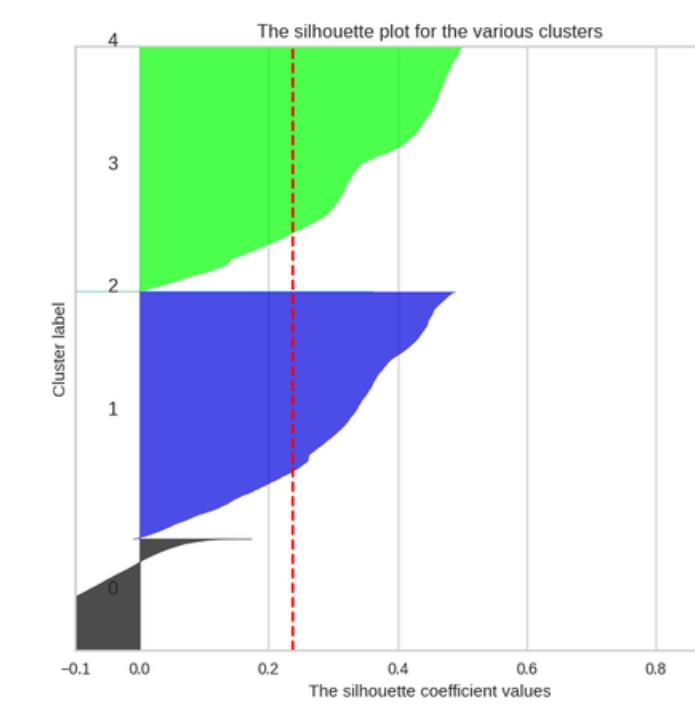
SEGMENTATION - CLUSTERING Y21



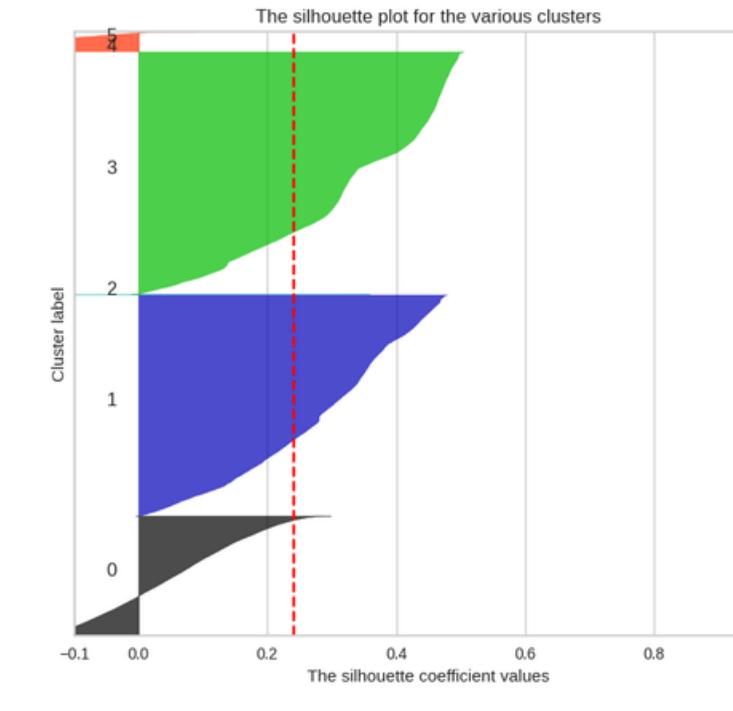
- There is no significant separation after K=4 although silhouette score improves slightly



K = 4



K = 5



K = 6

For `n_clusters` = 4 The average silhouette_score is: 0.23758185640280155
For `n_clusters` = 5 The average silhouette_score is: 0.23801813377458775
For `n_clusters` = 6 The average silhouette_score is: 0.24113597964305067

SEGMENTATION - CLASSIFICATION Y22

MODEL ACCURACY

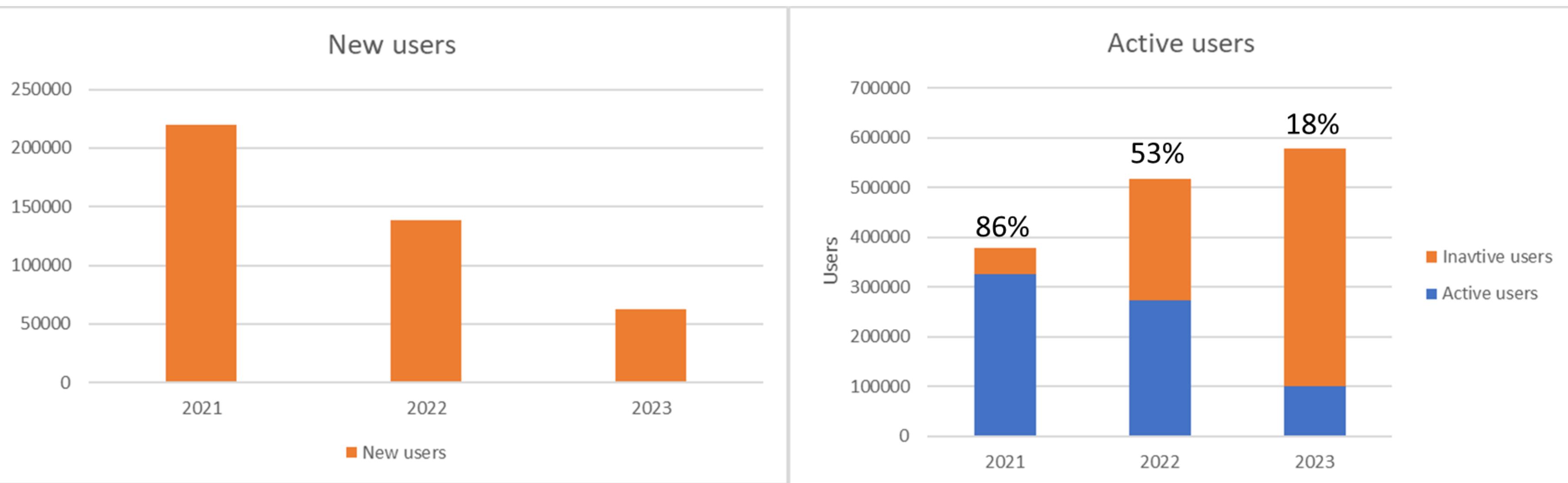
	Model	Accuracy	AUC	Recall	Prec.	F1
ridge	Ridge Classifier	0.7900	0.0000	0.7900	0.6452	0.7098
et	Extra Trees Classifier	0.7413	0.9793	0.7413	0.8198	0.6847
knn	K Neighbors Classifier	0.6585	0.8040	0.6585	0.6587	0.6583
lr	Logistic Regression	0.5583	0.7122	0.5583	0.5592	0.5472
rf	Random Forest Classifier	0.5232	0.9345	0.5232	0.5933	0.4375
nb	Naive Bayes	0.5226	0.7317	0.5226	0.5563	0.4366

DATA EXPLORE

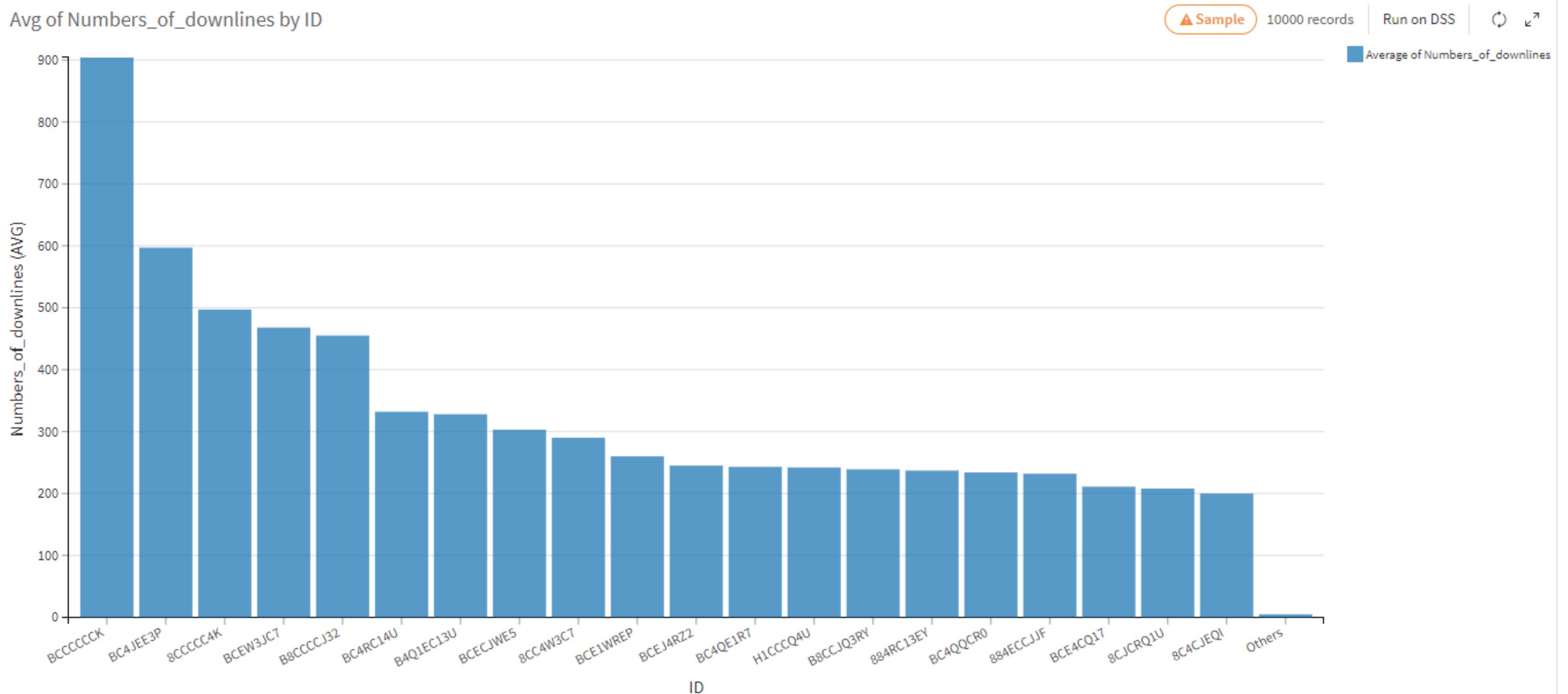
- GRAPH



GRAPH

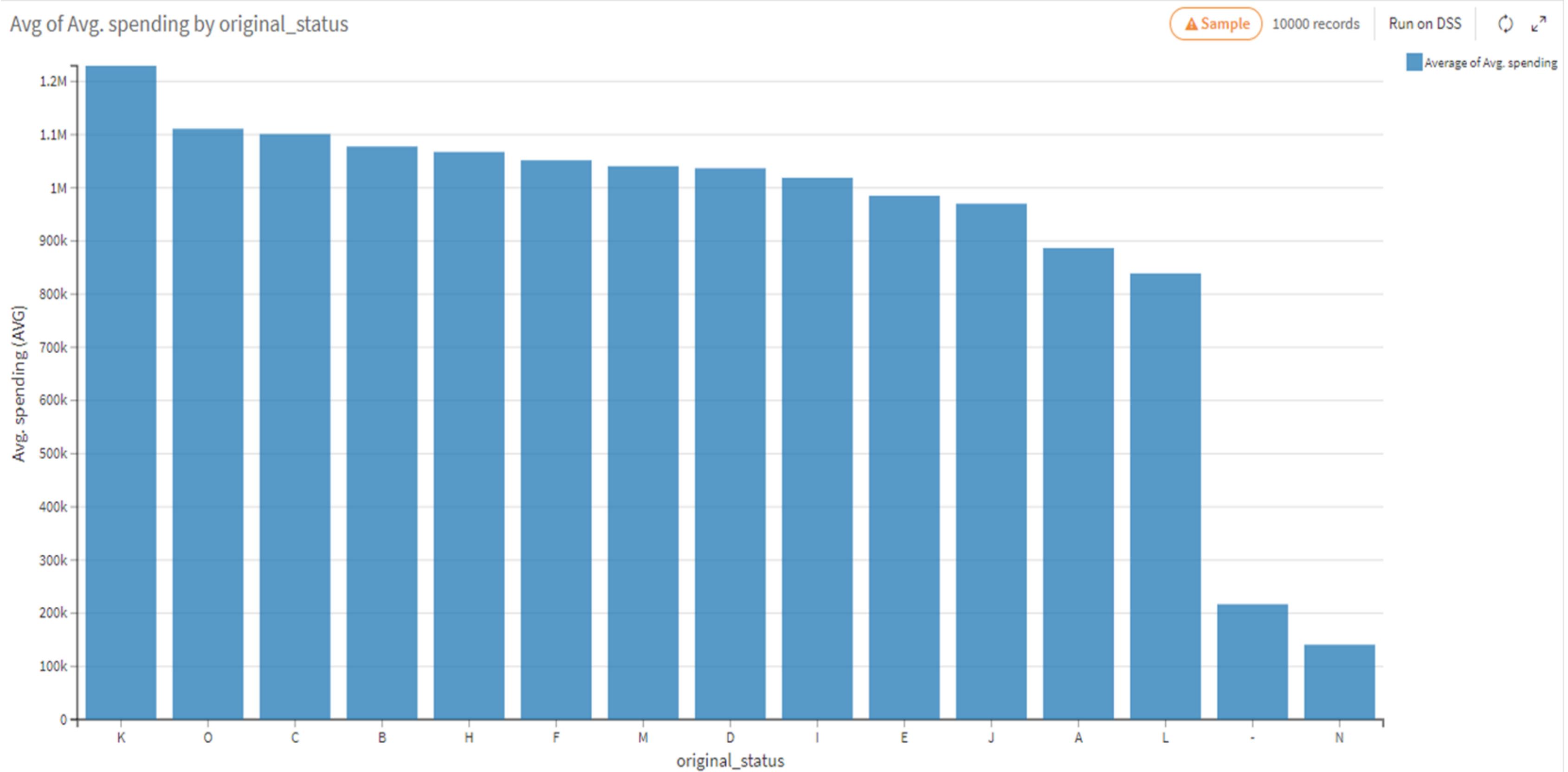


GRAPH



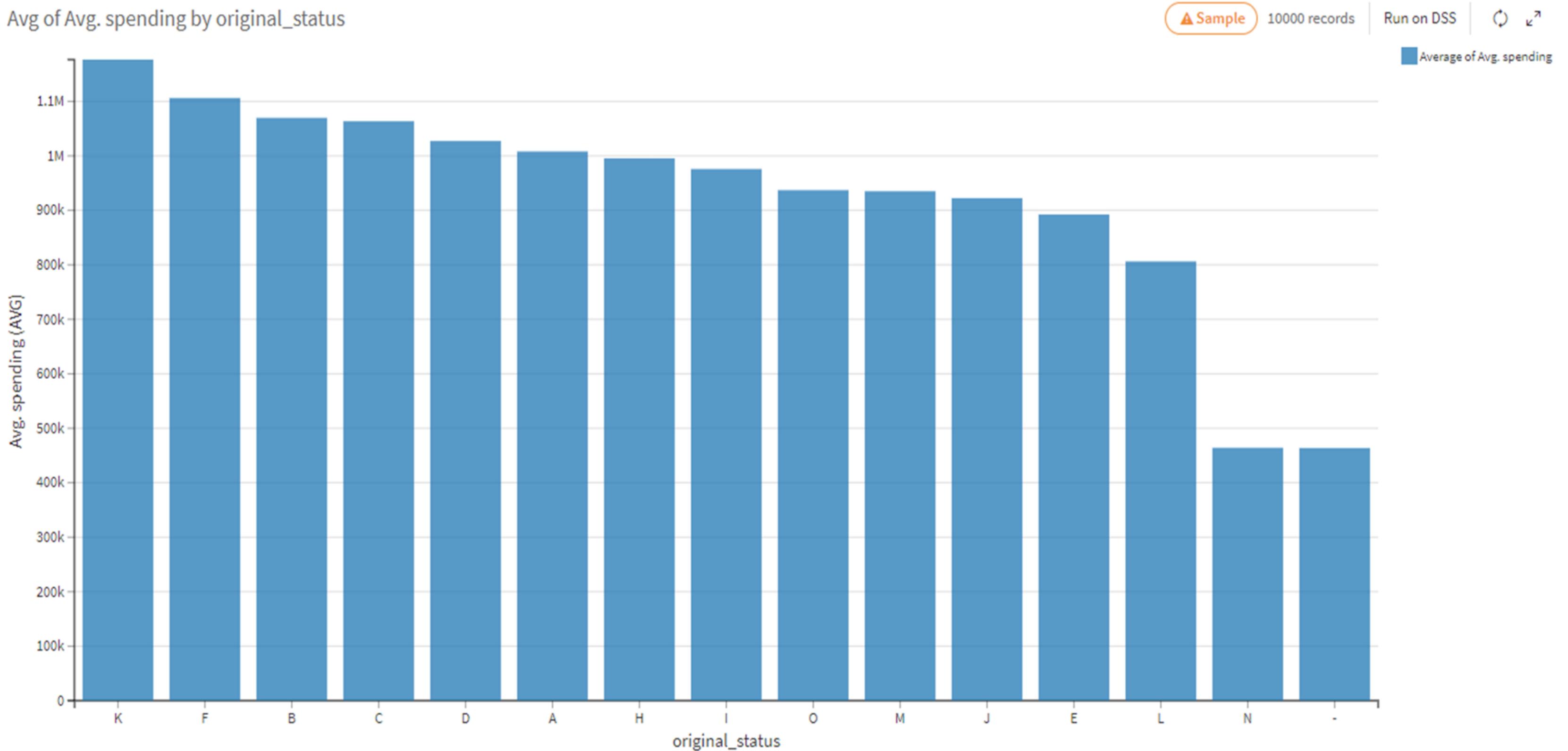
2021

Avg of Avg. spending by original_status



2022

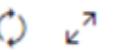
Avg of Avg. spending by original_status



⚠ Sample

10000 records

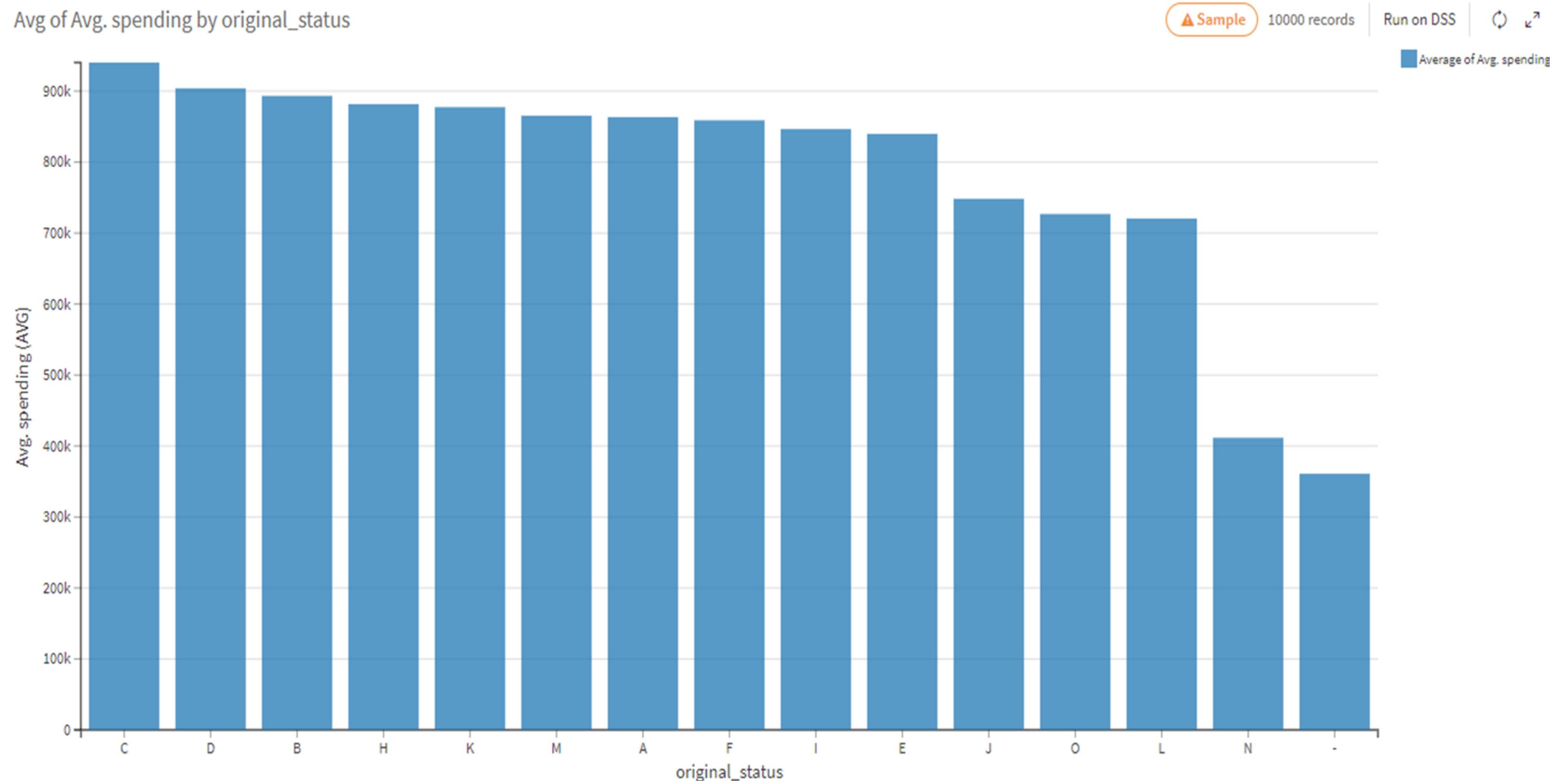
Run on DSS



Average of Avg. spending

2023

Avg of Avg. spending by original_status



⚠ Sample

10000 records

Run on DSS



Average of Avg. spending



THANK YOU

