

PROJECT / 머신러닝

2.RFM 분석 및 클러스터링 기반 고객 세분화를 통한 세그먼트 활용 전략 및 맞춤 상품 제안

프로젝트 개요

기획 의도	기존의 일반적인 상품 추천보다 고객의 특성과 선호도를 고려한 개인화된 상품 추천 서비스를 제공하여 고객 만족도를 높이는 것이 목표
문제 정의	RFM 분석과 K-means 클러스터링을 통해 고도화된 고객 세분화를 통해, 고객 특성에 맞는 맞춤 상품 제안
분석 목표	고객 세분화 및 상품 추천 결과를 기반으로 한 효과적인 전략을 제시하여 , 매출 증대 및 이탈 방지
Language	Python
Tool	Pandas

프로세스



데이터

데이터 설명	
데이터 이름	E-commerce
데이터 출처	Kaggle
데이터 사이즈	541909 rows X 8columns

데이터 전처리

데이터 전처리	
결측치	CustomerID는 25% 결측치 존재하지만,임의 대체 시 분석에 있어 편향이 발생할 수 있으므로, 삭제 처리
중복값	중복 데이터 삭제

RFM

RFM 계산

Recency

최근 거래 기준일 - 고객별 최근 구매 날짜

Frequency

고객별 구매 빈도 수

Monetary

고객 별 총 구매 금액

RFM 모형

Scoring 기법 : RFM 요인을 각각 5등급 간격으로 분류

Recency_Score	최근 구매 날짜일수록,높은 점수 부여
Frequency_Score	구매 횟수가 클수록,높은 점수 부여
Monetary_Score	구매 총액이 클수록,높은 점수 부여
RFM_Score	RFM 요소에 대한 점수를 각각 100,10,1의 자릿수로 사용하여 하나의 점수로 통합 ⇒ RFM 세그먼트에 사용

RFM 세그먼트

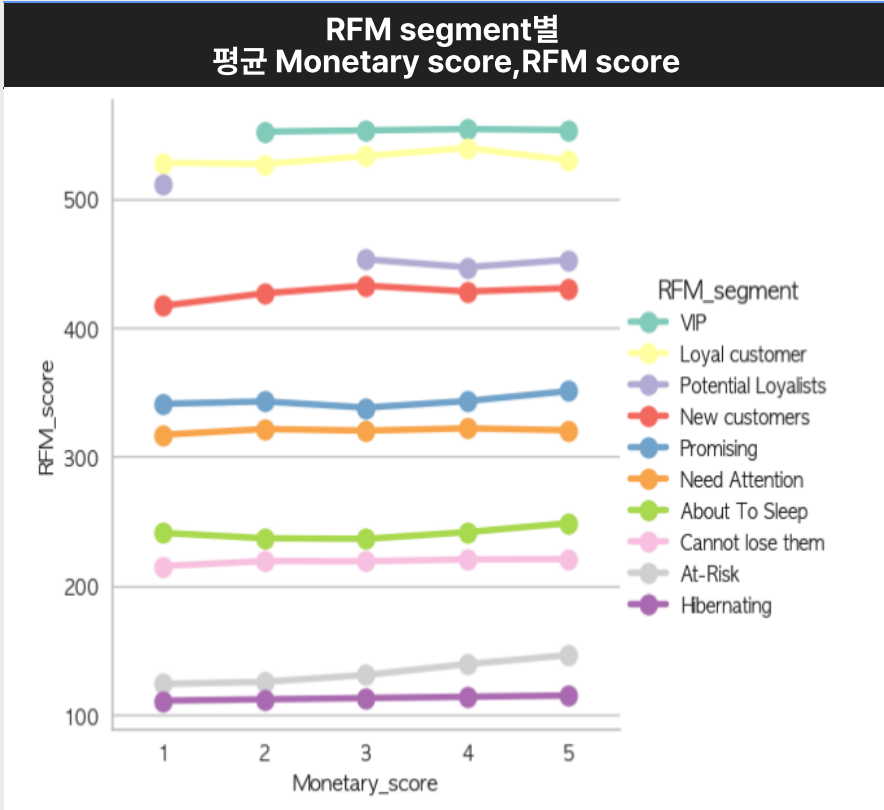
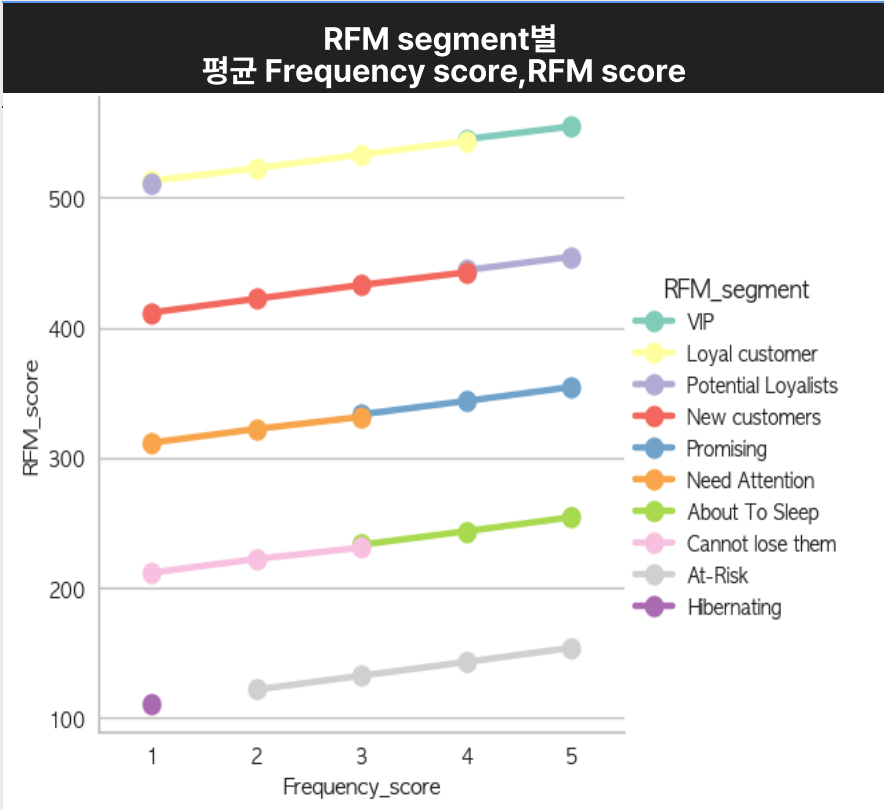
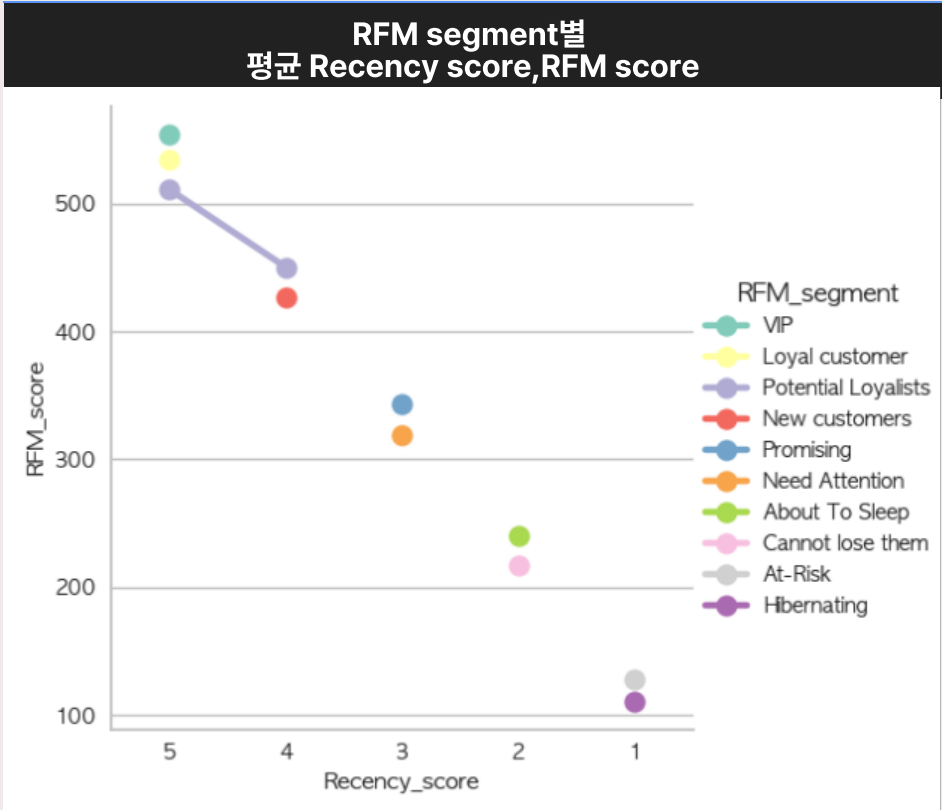
RFM_score	
>=545	VIP
>=511 and <545	Loyal custoemr
>=443 and <511	Potential Loyalists
>=411 and <443	New customers
>=333 and <411	Promising
>=311 and <333	Need Attention
>=231 and <311	About To Sleep
>=188 and <231	Cannot lose them
>=121 abd <188	At-Risk
>=110 and <121	Hibernating

RFM 세그먼트

PROJECT / 머신러닝

2.RFM 분석 및 클러스터링 기반 고객 세분화를 통한 세그먼트 활용 전략 및 맞춤 상품 제안

RFM 세그먼트 분석



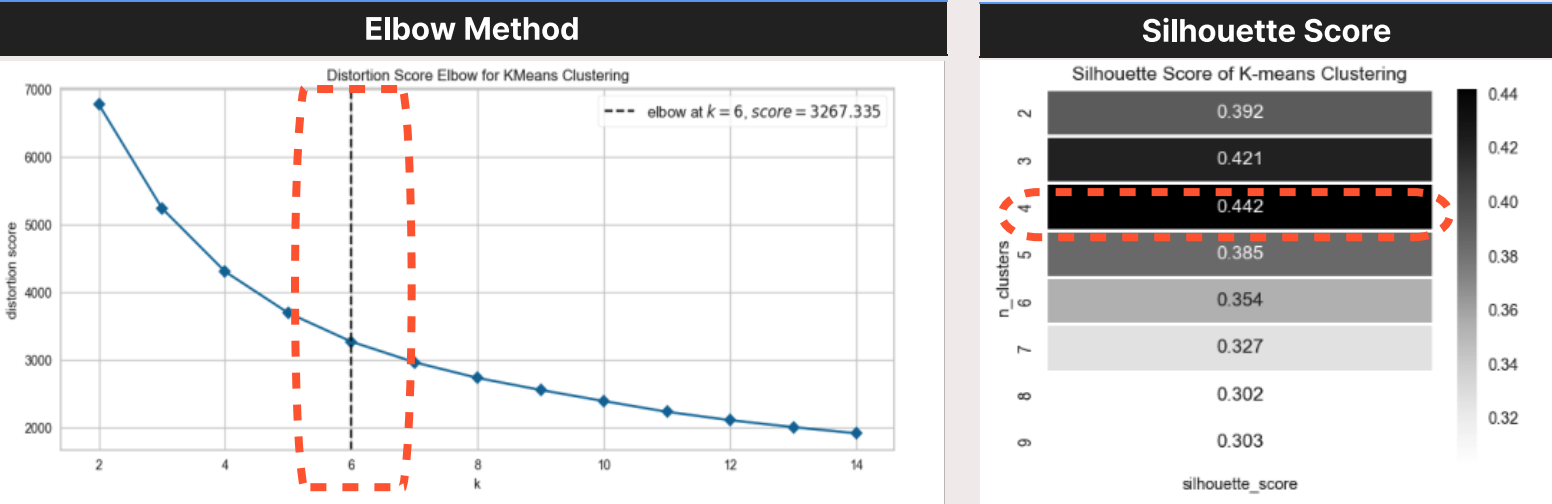
VIP	가장 최근 활동/ 가장 높은 구매 빈도 /가장 높은 구매금액 (높은 가치 창출 고객)
Loyal custoemr	최근 활동 /VIP에 비해 낮은 구매 빈도/ VIP에 비해 낮은 금액 소진 (충성도 높은 고객)
Potential Loyalists	상대적으로 최근 구매 /높은 구매 빈도/ 높은 구매 금액 (잠재적 고객)
New customers	최근 유입/ 낮은 빈도/낮은 구매 금액 (신규 고객)
Promising	평균적인 최근성/평균 빈도/평균 구매 금액 (가치를 높일 가능성 있는 고객)
Need Attention	낮은 구매 빈도/낮은 구매 금액 (관심이 필요한 고객)
About To Sleep	활동은 있었지만, 더 이상 활발하지 않은 고객 (잠들기 직전 고객)
Cannot lose them	평균 이하 최근성/낮은 구매 빈도/평균 구매 금액 값을 가진 고객(잃을수 없는 고객)
At-Risk	낮은 최근성/낮은 구매 빈도/평균 구매 금액 값을 가진 고객 (이탈 위험 고객)
Hibernating	가장 낮은 최근성/가장 낮은 빈도를 가진 고객

PROJECT / 머신러닝

2.RFM 분석 및 클러스터링 기반 고객 세분화를 통한 세그먼트 활용 전략 및 맞춤 상품 제안

고객 세그먼트

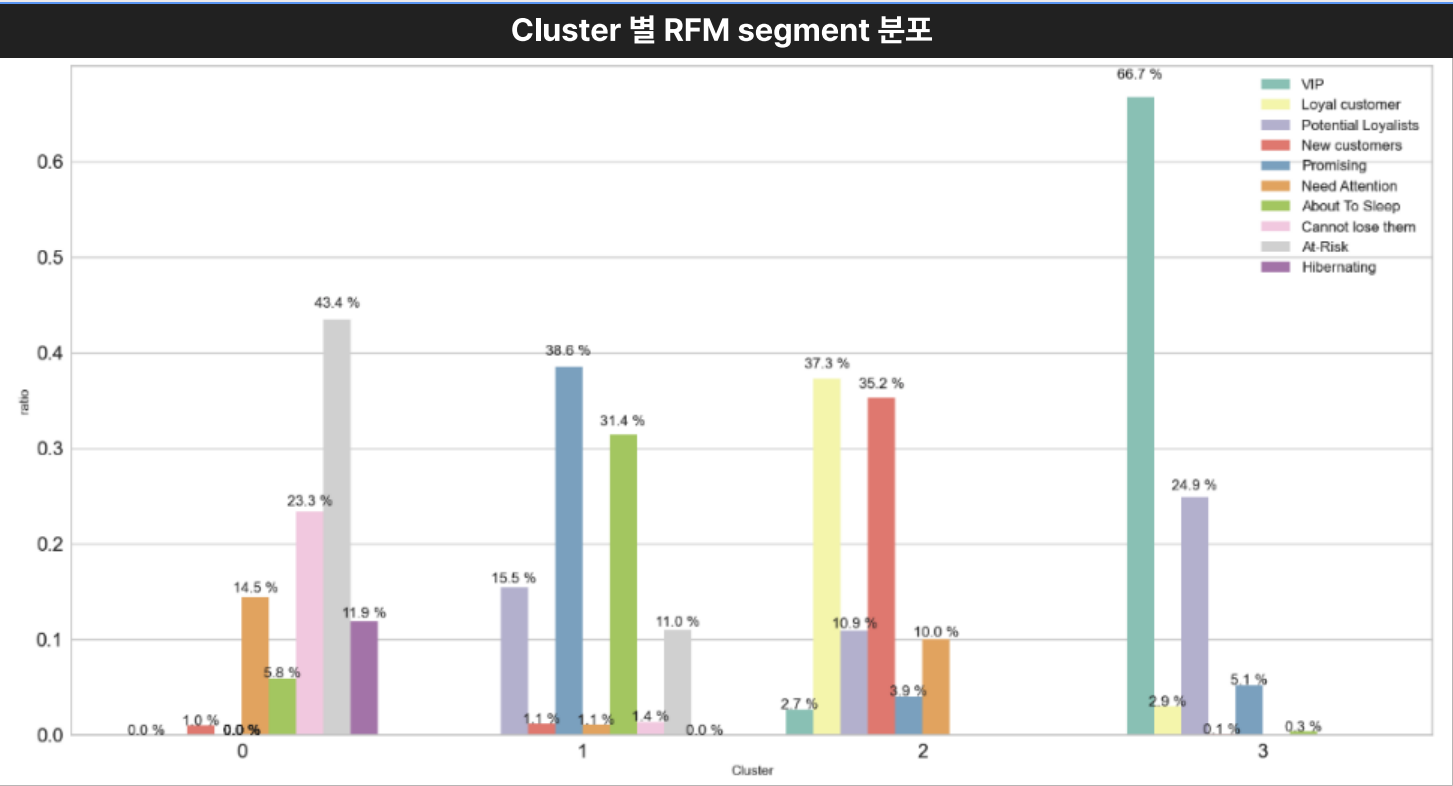
K-means 클러스터링



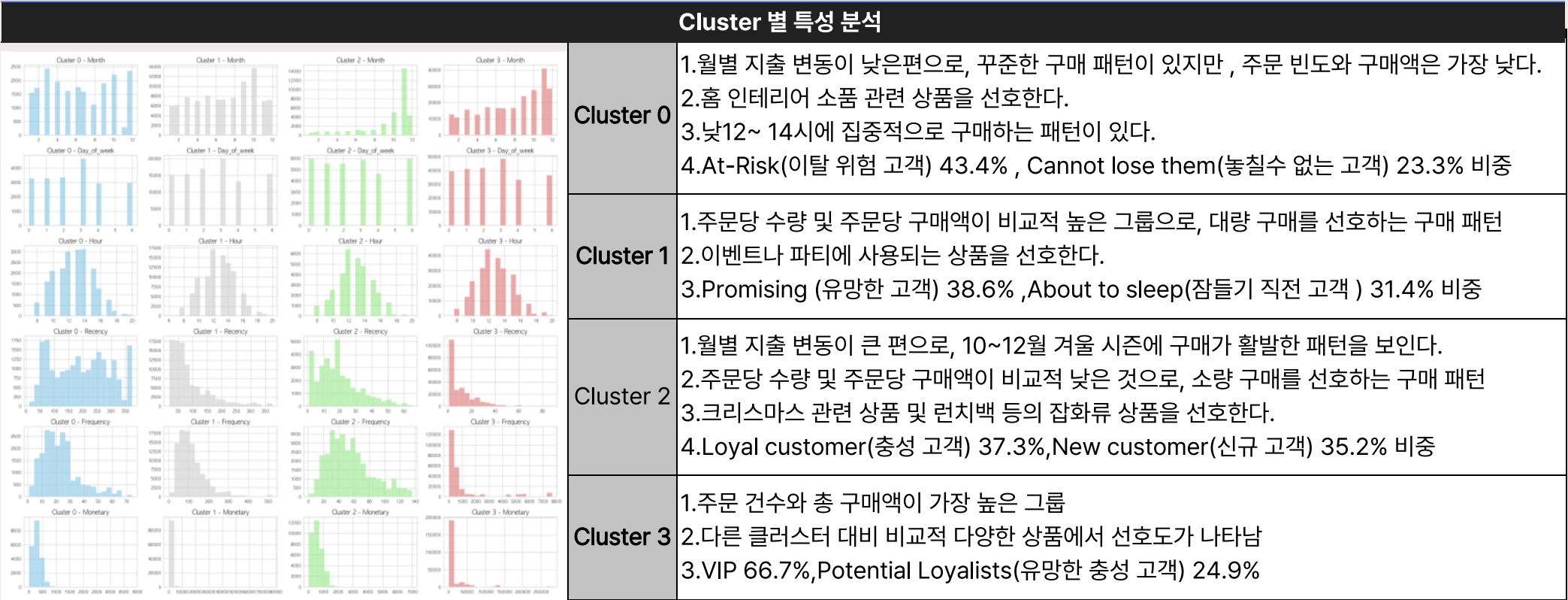
Elbow Method : 6 / Silhoutette Score: 4

두 지표를 함께 고려하여, 클러스터 개수 4 선택
⇒ 클러스터 4개로 분류했을때, 실루엣 계수는 0.442로 가장 높으며, 개별 군집의 평균 실루엣 계수 값이 비교적 균일하다.

클러스터 별 세그먼트 분포



클러스터 별 특성



PROJECT / 머신러닝

2.RFM 분석 및 클러스터링 기반 고객 세분화를 통한 세그먼트 활용 전략 및 맞춤 상품 제안

고객 세그먼트

BASKET Analysis

Cluster 0										
	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
17	(ROSES REGENCY TEACUP AND SAUCER, PINK REGENC...	(GREEN REGENCY TEACUP AND SAUCER)	0.022545	0.043636	0.021091	0.935484	21.438172	0.020107	14.823636	0.975344
3	(PINK REGENCY TEACUP AND SAUCER)	(GREEN REGENCY TEACUP AND SAUCER)	0.032000	0.043636	0.027636	0.863636	19.791667	0.026240	7.013333	0.900861
1	(PAINTED METAL PEARS ASSORTED)	(ASSORTED COLOUR BIRD ORNAMENT)	0.029091	0.061818	0.023273	0.800000	12.941176	0.021474	4.690909	0.950375
18	(GREEN REGENCY TEACUP AND SAUCER, PINK REGENCY...	(ROSES REGENCY TEACUP AND SAUCER)	0.027636	0.041455	0.021091	0.763158	18.409511	0.019945	4.047192	0.972558
6	(ROSES REGENCY TEACUP AND SAUCER)	(GREEN REGENCY TEACUP AND SAUCER)	0.041455	0.043636	0.029818	0.719298	16.483918	0.020009	3.407045	0.979959

Cluster 1										
	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
238	(SET/10 BLUE POLKADOT PARTY CANDLES)	(SET/10 PINK POLKADOT PARTY CANDLES)	0.023837	0.030647	0.021566	0.904762	29.522046	0.020836	10.178207	0.909718
239	(SET/10 PINK POLKADOT PARTY CANDLES)	(SET/10 BLUE POLKADOT PARTY CANDLES)	0.030647	0.023837	0.021566	0.703704	29.522046	0.020836	3.294552	0.996672
101	(HAPPY STENCIL CRAFT)	(MONSTERS STENCIL CRAFT)	0.035187	0.030647	0.023837	0.677419	22.103943	0.022758	3.004994	0.989580
100	(MONSTERS STENCIL CRAFT)	(HAPPY STENCIL CRAFT)	0.030647	0.035187	0.023837	0.777778	22.103943	0.022758	4.341657	0.984945
63	(GLASS APOTHECARY BOTTLE PERFUME)	(GLASS APOTHECARY BOTTLE TONIC)	0.031782	0.031782	0.021566	0.678571	21.350765	0.020556	3.012234	0.984451

Cluster 3										
	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
812	(LUNCH BAG SPACEBOY DESIGN , JUMBO BAG APPLES)	(LUNCH BAG RED RETROSPOT)	0.025404	0.135104	0.023095	0.909091	6.728827	0.019662	9.513857	0.873578
960	(LUNCH BAG SPACEBOY DESIGN , WHITE HANGING HEA...	(LUNCH BAG RED RETROSPOT)	0.024249	0.135104	0.020785	0.857143	6.344322	0.017509	6.054273	0.863314
794	(JUMBO BAG APPLES, LUNCH BAG BLACK SKULL.)	(LUNCH BAG RED RETROSPOT)	0.027714	0.135104	0.023095	0.833333	6.168091	0.019350	5.189376	0.861758
769	(JUMBO BAG PINK POLKADOT, LUNCH BAG WOODLAND)	(JUMBO BAG APPLES)	0.025404	0.080831	0.020785	0.818182	10.122078	0.018732	5.055427	0.924697
824	(JUMBO BAG PINK POLKADOT, JUMBO BAG OWLS)	(JUMBO BAG RED RETROSPOT)	0.027714	0.139723	0.021940	0.791667	5.665978	0.018068	4.129330	0.846981

Cluster 4										
	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
203	(RED HANGING HEART T-LIGHT HOLDER)	(WHITE HANGING HEART T-LIGHT HOLDER)	0.052023	0.167630	0.037159	0.714286	4.261084	0.028439	2.913295	0.807317
26	(CANDLEHOLDER PINK HANGING HEART)	(WHITE HANGING HEART T-LIGHT HOLDER)	0.041288	0.167630	0.028076	0.680000	4.056552	0.021155	2.601156	0.785935
55	(PINK REGENCY TEACUP AND SAUCER)	(GREEN REGENCY TEACUP AND SAUCER)	0.049546	0.061107	0.033031	0.686667	10.909910	0.030003	2.816680	0.955691
206	(REGENCY TEA PLATE GREEN)	(REGENCY CAKESTAND 3 TIER)	0.035508	0.156069	0.021470	0.604651	3.874246	0.015928	2.134648	0.769198
29	(CHARLOTTE BAG PINK POLKADOT)	(RED RETROSPOT CHARLOTTE BAG)	0.034682	0.049546	0.020644	0.595238	12.013689	0.018926	2.348181	0.949701

지지도 0.2이상,신뢰도 0.7이상 연관규칙 5가지 도출

⇒ 두 제품 간의 연관성이 높으므로, 이를 함께 마케팅하거나 특별한 할인 혜택을 제공할 수 있다.
⇒이를 함께 판매하는등의 상품 번들구성을 기획할 수 있다.

Recommend System

클러스터 기반 추천 시스템													
Description	4 PURPLE FLOCK DINNER CANDLES	50'S CHRISTMAS GIFT BAG LARGE	DOLLY GIRL BEAKER	I LOVE LONDON MINI BACKPACK	OVAl WALL MIRROR DIAMANTE	RED SPOT GIFT BAG LARGE	SET 2 TEA TOWELS I LOVE LONDON	SPACEBOY BABY GIFT SET	10 COLOUR SPACEBOY PEN	12 COLOURED PARTY BALLOONS	... ZINC HEARTS PLANT POT HOLDER	ZINC HERB GARDEN CONTAINER	
CustomerID													
12350.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
12353.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
12355.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
12361.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
12363.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
...
18276.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
18277.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
18278.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
18280.0	0.0	0	0	0	0	0	0.0	0	0	0	...	0	0
18281.0	0.0	0	0	0	0	0	0.0	1	0	0	...	0	0

예시: 클러스터1에서 사용자12358.0 10개의 상품을 추천

cluster_id = 1
user_id = 12358.0
num_recommendations = 10
recommended_items_cluster_0 = recommend_items_in_cluster(cluster_id, user_id, num_recommendations)
recommended_items_cluster_0

recommend_items

0 EGG FRYING PAN MINT
1 EGG FRYING PAN IVORY
2 CHARLIE & LOLA WASTEPAPER BIN BLUE
3 EGG CUP NATURAL CHICKEN
4 CHARLIE & LOLA WASTEPAPER BIN FLORA
5 EGG FRYING PAN BLUE
6 CERAMIC CAKE BOWL + HANGING CAKES
7 CERAMIC HEART FAIRY CAKE MONEY BANK
8 FANCY FONTS BIRTHDAY WRAP
9 CHAMPAGNE TRAY BLANK CARD

클러스터 기반 추천 시스템 소개

사용자를 클러스터로 그룹화하여 취향 파악 ⇒ 각 클러스터에 대한 추천 제공
이미 구매한 상품을 제외하고, 클러스터와 고객 기반으로 n개의 상품을 추천

클러스터 별 활용 전략

Cluster 0	<p>1.At-Risk (이탈 위험 고객) 리워드 및 할인 제공: 추가 리워드나 할인을 제공하여 이탈 방지 설문조사 및 피드백 요청: 고객의 불만,이탈 원인을 파악하기 위해 설문조사요청⇒서비스나 상품 개선 시행.</p> <p>2.Cannot lose Them(잃을수 없는 고객) 개인화된 혜택 제공: 할인, 멤버십 혜택, 또는 독점적인 프로모션을 통해 이탈 방지 특별한 이벤트 및 선점 기회 제공:특별한 이벤트나 선점 기회를 제공하여 구매 유도 낮 12~14시에 특별한 프로모션 제공</p>	<p>1.홈 인테리어 관련 상품에 대한 개인화된 추천 서비스 제공. 2.이탈 위험이 있는 사용자에게는 이전에 관심을 보였던 상품, 카테고리를 중심으로 개인화된 상품 추천</p>
Cluster 1	<p>1.Promising(유망한 고객) 추가 혜택 및 프로모션 제공: 추가 혜택이나 독점적인 프로모션을 제공하여 더 많은 구매를 유도 추가 서비스 및 제품 소개:신규 서비스나 제품을 소개하여 구매 유도</p> <p>2.About To Sleep (곧 잠들 예정의 고객) 개인화된 알림 및 특별 제안 제공: 개인화된 알림이나 특별 제안을 통해 다시 관심을 유도 재구매 유도 프로그램 도입: 재구매를 유도하는 CRM 메세지 전송 및 리텐션 마케팅 진행 포인트 및 할인 유도 :추가 포인트나 할인을 제공하여 구매 유도</p>	<p>1.추가 구매를 유도하는 번들 상품 제안:장바구니에 담긴 상품 간의 연관성을 고려하여,번들 상품을 제안 2.개인화된 추천 시스템 : 이벤트 상품을 구매하는 경향으로 관련 상품 추천</p>
Cluster 2	<p>1.Loyal Customer (충성 고객) 추가 혜택 및 리워드 프로그램: 충성도를 인정하고 보상하는 리워드 프로그램을 제공하여 계속해서 구매 유도 개별적인 감사 메세지: 고객에게 개별적으로 감사의 뜻을 전하는 메시지 전송</p> <p>2.New Customers (신규 고객) 환영 혜택 및 할인 제공:초기 구매에 대한 환영 혜택이나 할인을 제공하여, 브랜드에 대한 긍정적인 경험을 제공 계절 별 프로모션 및 할인:겨울 시즌에 특화된 할인 프로그램 도입하여 계절에 따른 구매를 유도</p>	<p>1.시즌별 상품 추천 :선호도 분석을 통해 겨울 시즌에 가장 선호하는 상품 식별 및 계절 상품 추천</p>
Cluster 3	<p>1.VIP 고객 VIP 멤버십 프로그램 운영: VIP와 Potential Loyalists 그룹에 대한 멤버십 프로그램 도입. 선점적인 새로운 상품 소개:새로운 상품이나 서비스를 먼저 체험할 수 있는 기회를 제공하여 로열티 증대 우선 배송 및 서비스 제공:빠른 배송 서비스나 24시간 고객 서비스 등을 우선적으로 제공</p> <p>1.Potential Loyalists (잠재적 충성고객) 추가 혜택 및 가치 제시:높은 구매 빈도와 금액에 대한 추가 혜택을 제시하여, 실제 충성고객으로 성장하도록 유도 추가 서비스 및 소식 제공 : 최신 상품이나 이벤트에 대한 정보를 제공하여, 고객의 참여를 유도</p>	<p>1.최근 구매 이력 기반 추천:VIP 및 Loyal Customer의 최근 구매 이력을 기반으로 ,유사한 상품 추천 2.고가 상품 및 프리미엄 아이템 추천:높은 구매 빈도와 금액을 갖는 이 세그먼트에게는 고가 상품이나,프리미엄 아이템을 중심으로 추천</p>