

KISA-신한카드 데이터톤

가구원수 예측모델과 연령에 따른 소비패턴 분석

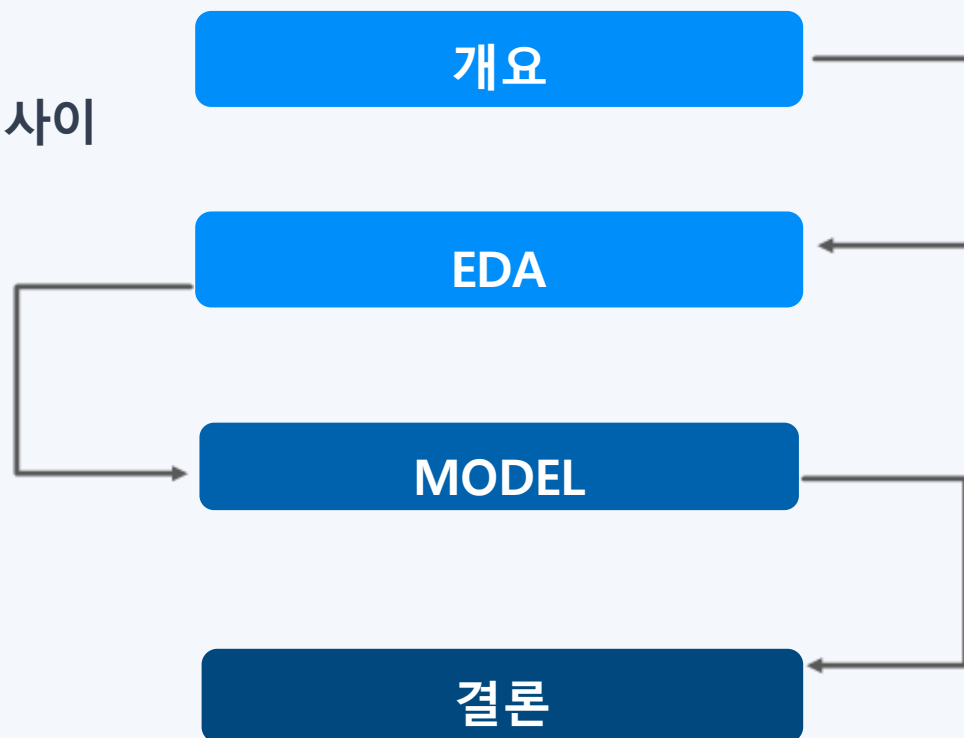
📍사당역 4번출구

김정의, 이원도, 류경희

목차

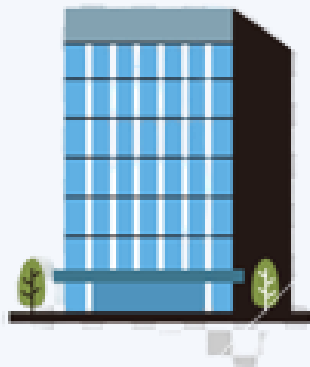
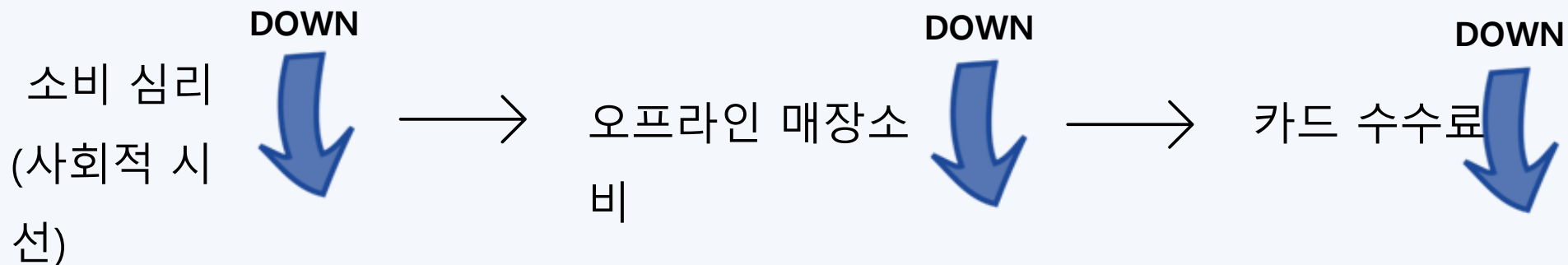
GOAL

가구 형태를 예측하는 최적의 모델 찾기 & 인사이트 도출



1. 개요

✓ 포스트 코로나 시대, 언택트 소비가 증가하다

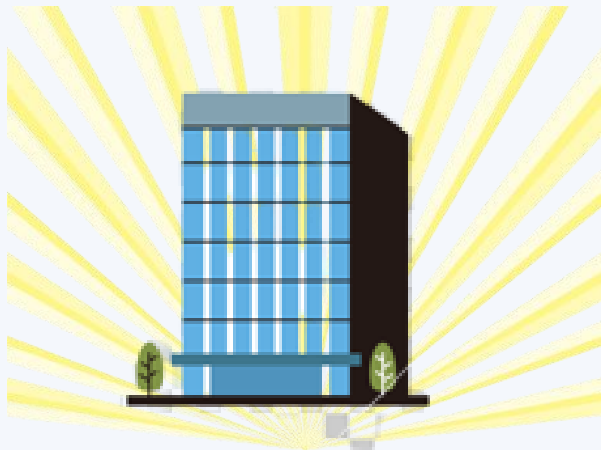


2020년 1분기 기준,
신용카드 수수료 수익 전년 동기보다 **40.7% 감소** (510
억원)

1. 개요

✓ 포스트 코로나 시대, 언택트 소비가 증가하다

 ShinhanCard



<상반기 당기순이익>

2713억



2019년 상반기

익>



3025억



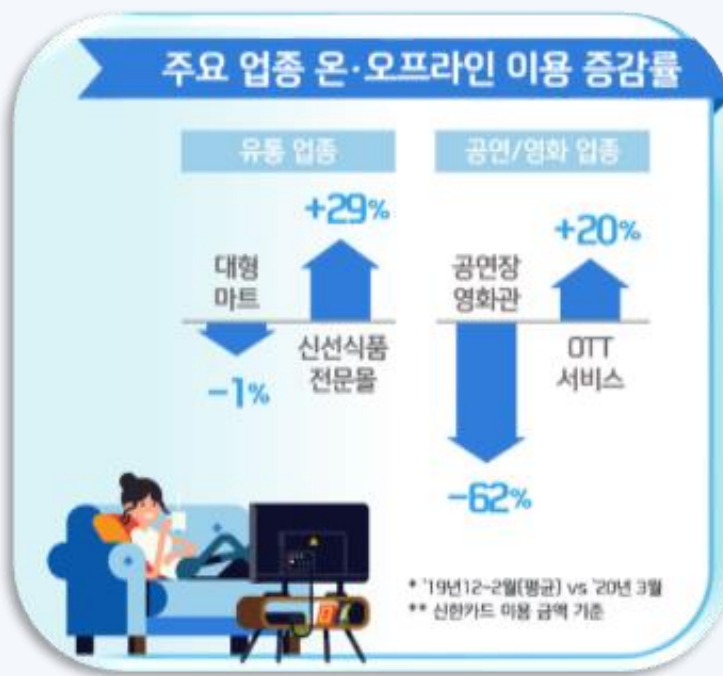
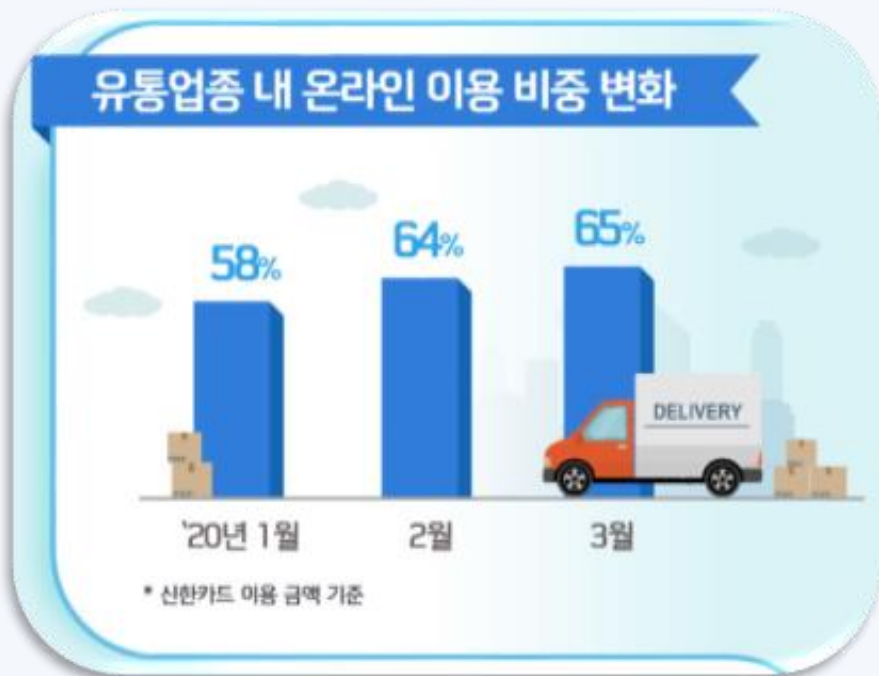
2020년 상반기

Why? -VAN 사업자에게 주는 수수료가 감소 (VAN수수료 온라인 < 오프라인)

대출(현금서비스, 카드론)에 의한 수익 증가

1. 개요

✓ 가구원수만이 아닌 가구원수와,나이를 고려한 세분화를 통한 인사이트 도출하기

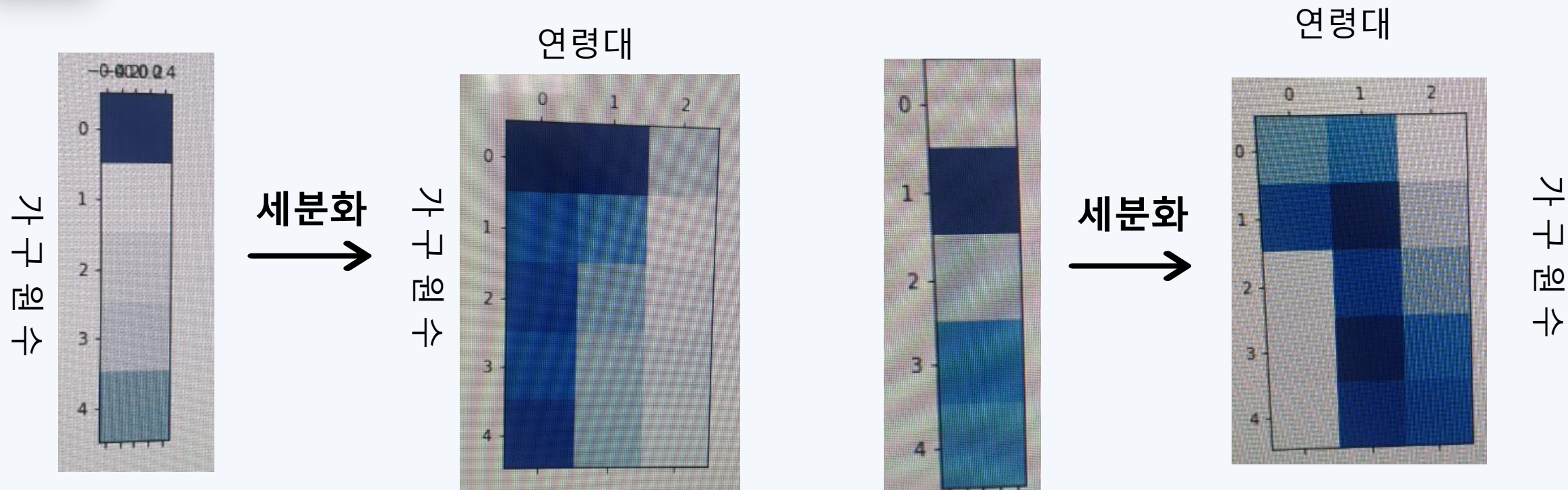


언택트시대 소비트렌드

- S** - 소비 중심 온라인으로 스위칭
- H** - 집 주변서 소비하는 홈라이프
- O** - 건강 위생 힐링 소비 급성장
- C** - 소비기간 구매 등 패턴 변화
- K** - 시니어 중심 디지털 소비 확대

2. EDA

1)나이를 고려한 세분화의 필요성



'편의점' 지출액에서

20대 1인가구와 50대이상 1인가구의 지출액
다름

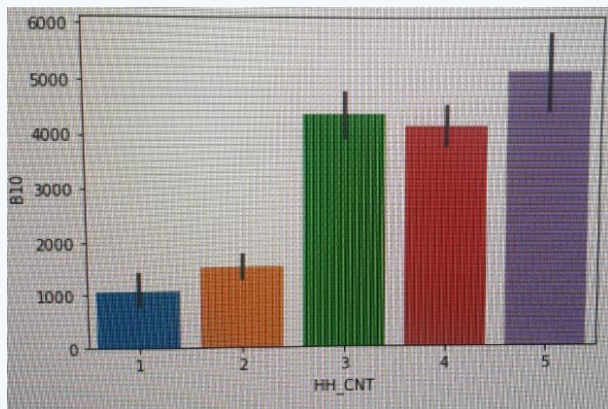
'항공/여행' 지출액에서

연령대에따라 소비패턴 상당히 다름

= 가구원수가 같다고 하더라도, 나이에 따라 전혀 다른 소비 패턴을 보

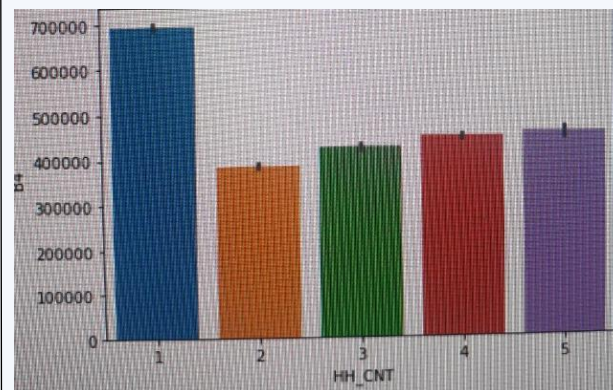
2. EDA

2) 모델링에 필요한 주요 변수



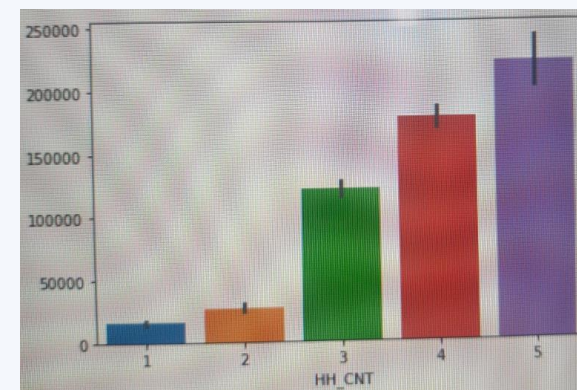
아동, 유아
복

1,2인 가구와 3,4,5인가구의 차이



편의점

1인, 2인가구 차이

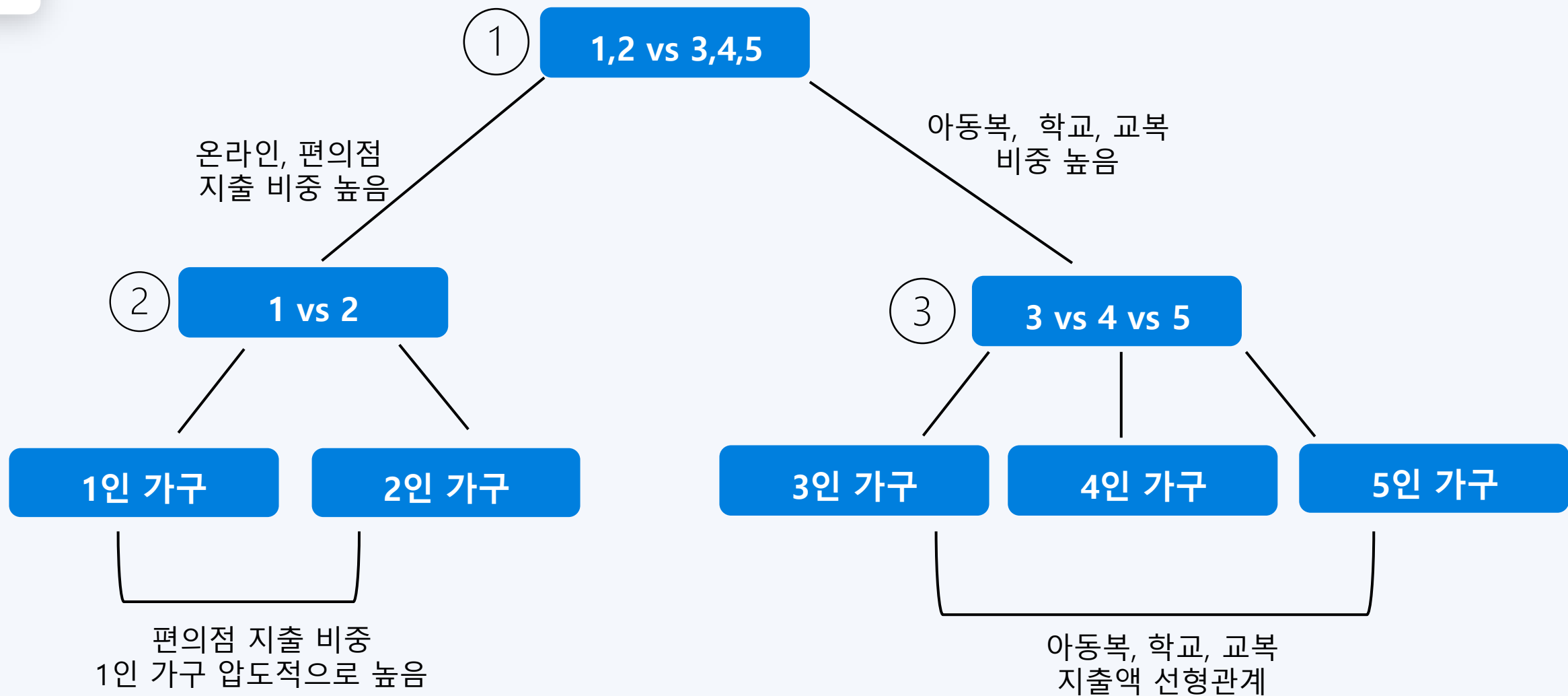


학교

3,4,5인 가구 => 선형성

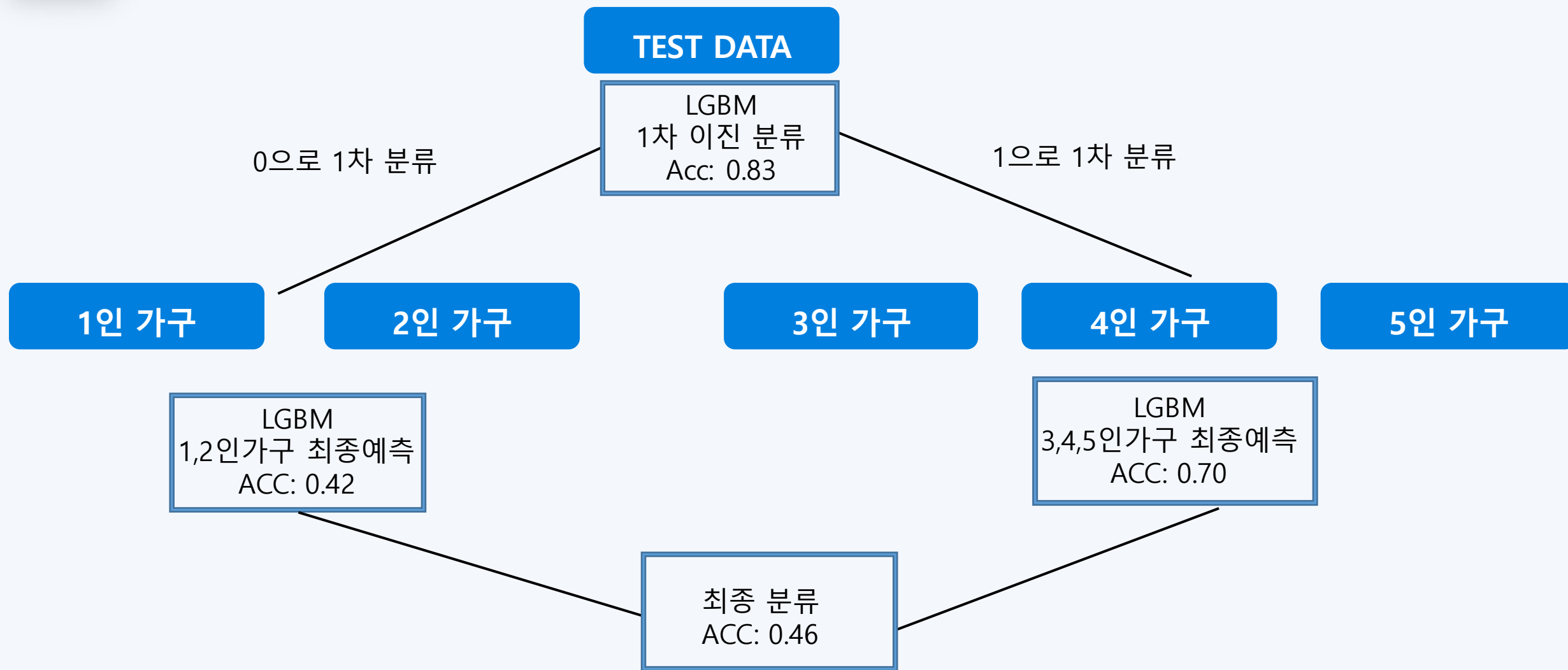
3. Model

✓ 모델링을 위한 주요 변수?



3. Model

✓ 모델링을 위한 주요 변수?



3. Model

✓ 모델의 정확도

		예측				
		1	2	3	4	5
실제	1	11643	4851	1971	2127	16
	2	3548	9947	3366	3212	16
	3	1820	5641	5869	8222	44
	4	1045	2309	4160	14599	90
	5	230	629	826	3797	44

	Accuracy	Micro F1	Macro F1
Test Data	0.468	0.386	0.468

4. 결론

✓ 가구원수를 추정을 통한 인사이트 도출하기



가구원수 예측



‘맞춤 마케팅’

Ex) 가구원수는 4인으로 예상되지만 신한카드를 혼자 사용하는 것으로 예상되는 경우

4. 결론

✓ 세분화된 전략



1인 가구 20대 남자

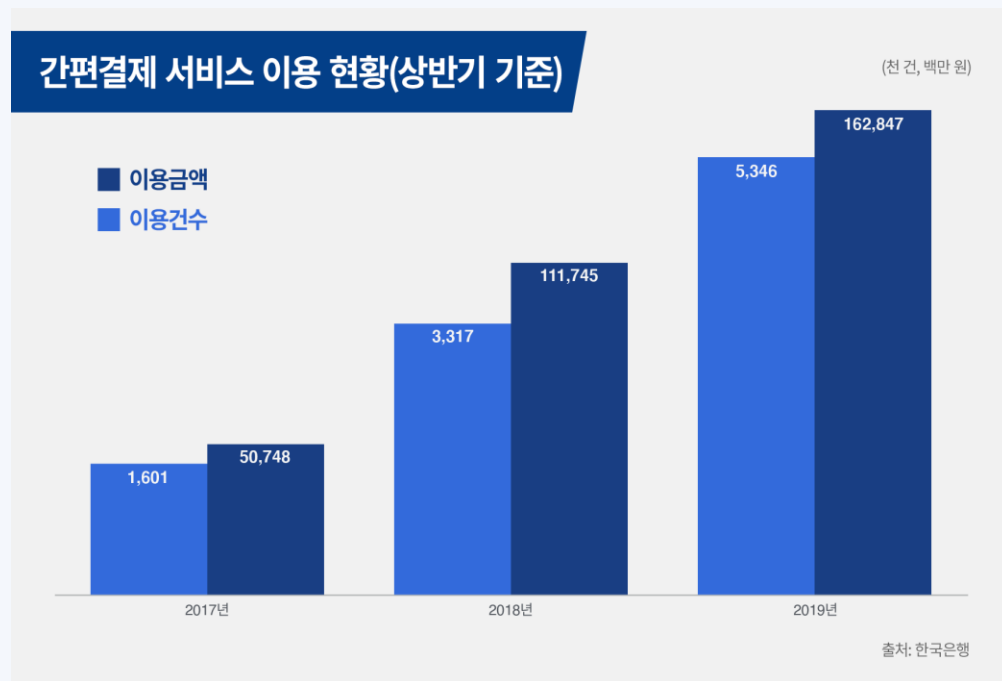


4인 가구 20대 남자

같은 연령, 성별이라도 **가구원수와 연령으로 세분화 함**에 따라 다른 소비
패턴을 보인다

4. 결론

✓ 세분화된 전략



- ✓ 간편결제 서비스 이용건수 지속적 증가
- ✓ 온라인 소비 가속화되면서 중요성 증대



- ✓ 고객 소비 특성 개발, 소비 테마별 AI 추천 모델에 적용
- ✓ 시간, 장소, 상황에 맞는 식당, 카페, 마트 등 추천