



REAL

SUMMIT 2022

SAMSUNG SDS

데이터 결합, 어디서부터 시작해야 하는가?

삼성SDS 데이터결합분석센터 이지면 프로

AGENDA

- I. 개요 : 데이터 결합은 무엇인가
- II. 효과 : 어떤 시나리오가 가능한가
- III. 방법 : 어떻게 결합해야 하는가

I.

개요 : 데이터 결합은 무엇인가

숫자로 보는 국내 데이터 산업은?

2021년(추정) *직접 매출 규모

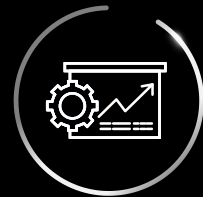
14조 9천억원

2020년 대비 **22%** 성장



데이터 처리 및 관리 솔루션 개발·공급업

• 3조 566억 원 (CAGR 14.2%)



데이터 구축 및 컨설팅 서비스업

• 5조 2,446억 원 (CAGR 13.5%)



데이터 판매 및 제공 서비스업

• 데이터 판매·중개 서비스 : 1.86조 (CAGR **28.0%**)

• 정보 제공 서비스 : 4.74조 (CAGR **32.7%**)

* 출처 : 2021 데이터 산업 백서 - K-Data * CAGR : 2016년-2021년

개인정보 활용이 가능? 데이터 3법 개정을 계기로...

2020.01.09 | 과학기술정보통신부 발표

인공지능 시대와 데이터 경제를 선도할 수 있는 제도적 기반이 마련되었습니다.

과학기술정보통신부는

데이터 3법 개정안의 국회 통과에 따른 후속조치로

4차 산업혁명 시대 핵심 자원인




데이터 개방·유통 확대를 추진하고, 데이터 간 융합과
활용 촉진을 통해 데이터 산업 육성을 본격 지원한다고 밝혔다.

이번 개정된 데이터3법은 **개인정보의 보호를 강화**하면서도

데이터 활용 활성화를 통한 관련 산업의 발전을 조화롭게
모색할 수 있도록 현행 제도를 보완하고 있다.

* 출처 : 대한민국 정책브리핑

개인정보의 개념은? 가명정보는 개인 동의 없이 활용 가능

	개념	예시	특징
실명 정보	특정 개인에 관한 정보로서 개인을 알아볼 수 있게 하는 정보	 홍길동 990909-1234567 010-1234-5678 서울 도봉구 방학로 4길 google@gmail.com 통신요금 7만 6550원, 폐렴	<ul style="list-style-type: none"> • 사전 동의 받은 경우에 한해 활용 가능 • 개인별로 차별화, 맞춤형 서비스 제공 가능
가명 정보	추가 정보의 사용 없이는 특정 개인을 알아볼 수 없게 조치한 정보 <small>*추가정보: 알고리즘, 맵핑데이터 등</small>	 홍** 99년생, 남자 서울 도봉구 *****@gmail.com 통신요금 7~8만원, 기관지질환	<ul style="list-style-type: none"> • 통계작성, 과학적 연구, 공익적 기록보존의 경우 개인 동의 없이 활용 가능 • 익명 정보 대비 활용 범위 넓음
익명 정보	더 이상 개인을 알아볼 수 없게(복원 불가능할 정도로) 조치한 정보	 - 20대, 남자 통신요금 7만 6500원, 기관지염	<ul style="list-style-type: none"> • 개인정보 보호법의 적용 대상이 아니므로 언제든지 동의 없이 활용 가능

* 출처 : 가명정보 처리 가이드라인, 개인정보보호위원회

데이터 3법에 대해 잠시 살펴보면... 가명정보 활용을 허용

법적 효력 없음

데이터 활용 위한 법적 근거 마련

민간 데이터 생산 거래 활용 촉진

개인정보 비식별 조치 가이드라인
(’16.6)

데이터 3법 개정
(’20.8)

데이터기본법 국회 통과
개인정보보호법 개정안 의결
(’21.9)

개정 前 개인정보 이용 시 사전에 동의를 받아야 함

개인정보보호법

- 가명정보 정의 및 활용 근거 마련
- 아래 목적 시 사전 동의 불필요



통계 작성



과학적 연구



공익적 기록

신용정보법

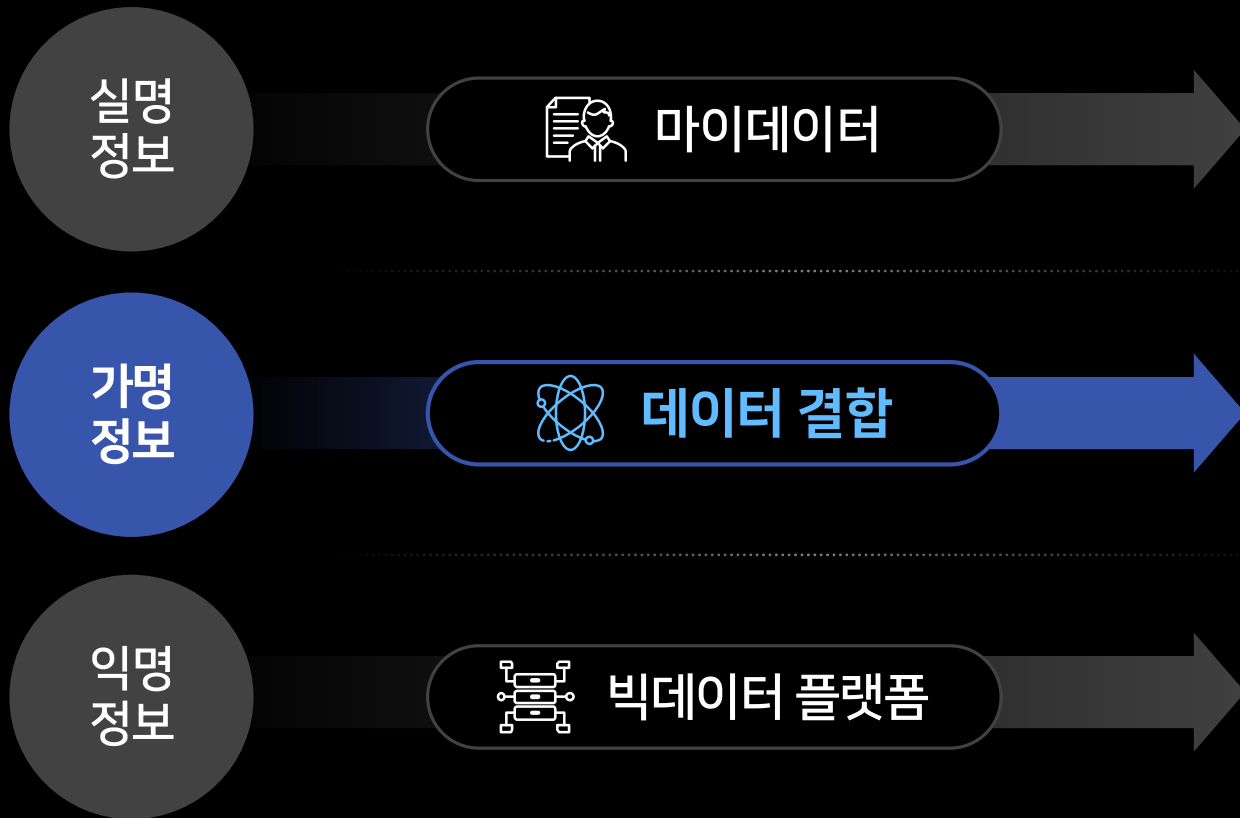
- 신용 정보의 가명 활용을 위한 근거 마련
- 금융 마이데이터 정의

**개인 동의하에
관련 금융 데이터 이동이 가능함**

정보통신망법

- 개인정보 규제 권한을
'개인정보보호법'으로 이관

가명정보의 활용은? 데이터 결합이라는 새로운 방식을 통해...



신규 제도·서비스 등장

- ✓ 분산된 개인 정보를
개인 동의 下 이동, 분석
- ✓ 각 기업·기관들의 개인 정보를
개인 동의 없이 가명처리, 결합
- ✓ 공공, 민간이 보유한 익명정보를
공유 개방

데이터 결합이란? 내가 보유하지 못한 고객 속성 데이터에 접근

“서로 다른 기업·기관이 보유한 **동일한 개인에 대한 정보**를 가명처리하고 결합하여 활용하는 것”

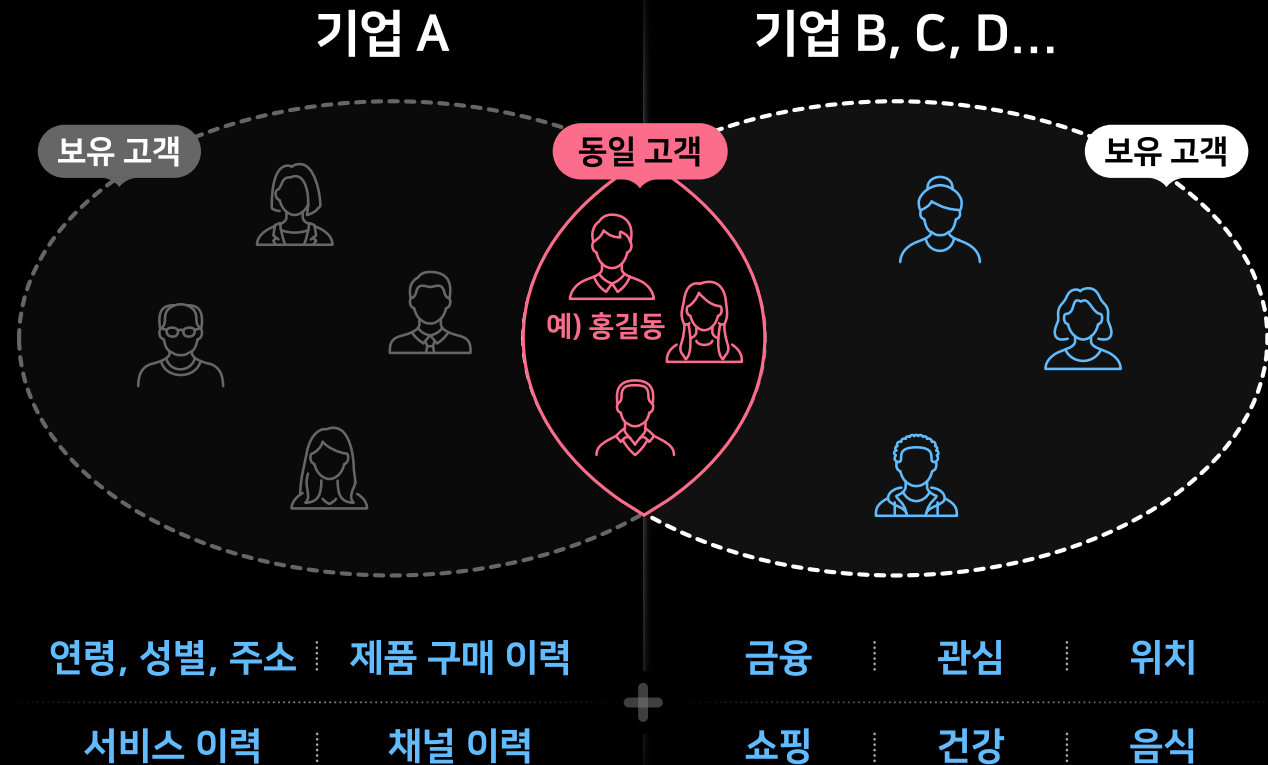
개인정보보호법 제28조의3 제1항

3대 목적

통계 작성

과학적 연구

공익적 기록

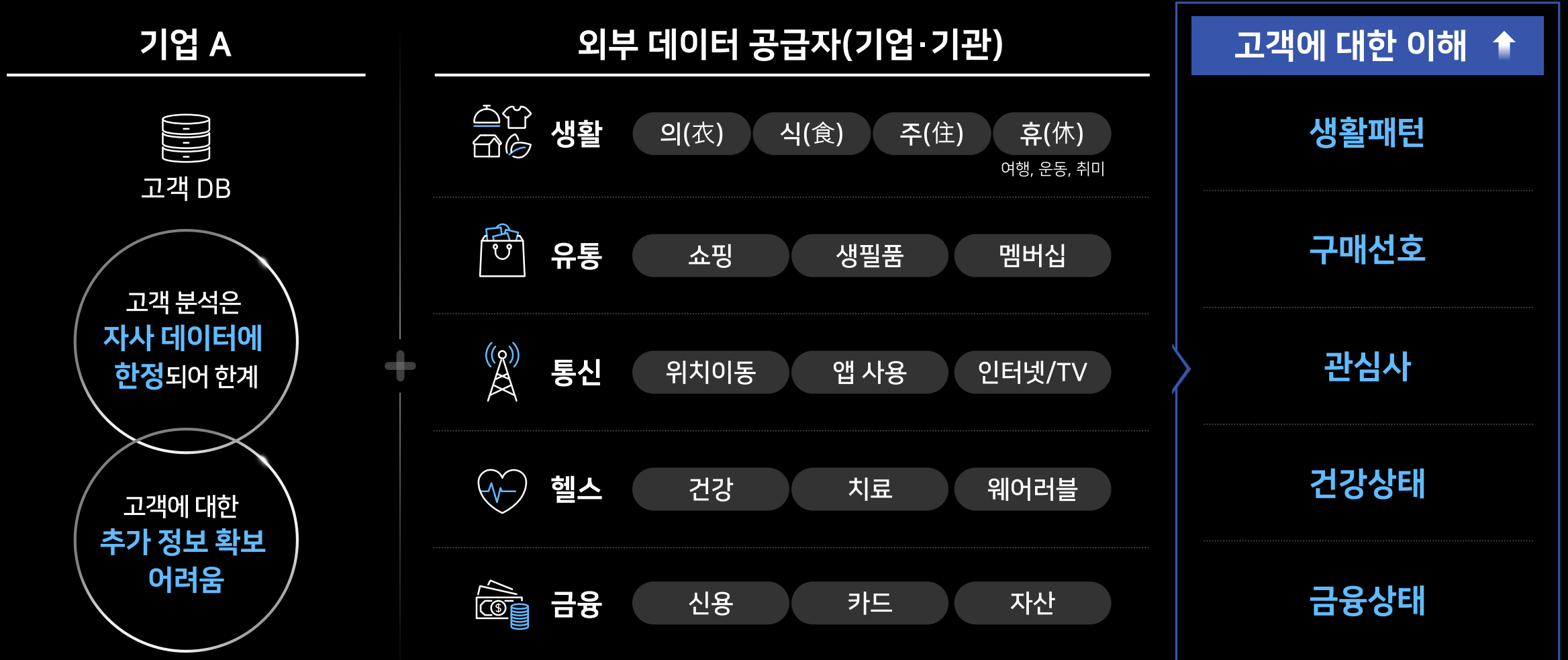


* 시장조사를 위한 통계 등 상업적 성격, 신기술·제품·서비스 개발 및 실증을 위한 산업적 연구도 해당

II.

효과 : 어떤 시나리오가 가능한가

결합을 통한 기대효과는? 고객 이해의 폭 확대, 새로운 인사이트 발견



사례 | 호텔체인

별도 법인으로 분리되어 있는 동일 브랜드 호텔을 방문한 고객정보를 분석하여 통합 멤버십 도입 효과를 예측해 볼 수 있었습니다.

Challenge

- ✓ COVID-19로 매출 부진 심화
- ✓ 통합 멤버십 제도 신설을 검토
- ✓ 그러나 각 호텔은 독립법인으로, 고객 정보의 통합 분석이 불가능

통합멤버십 적용하려면 비용이 많이 드는데 맞는 결정인지 모르겠어요

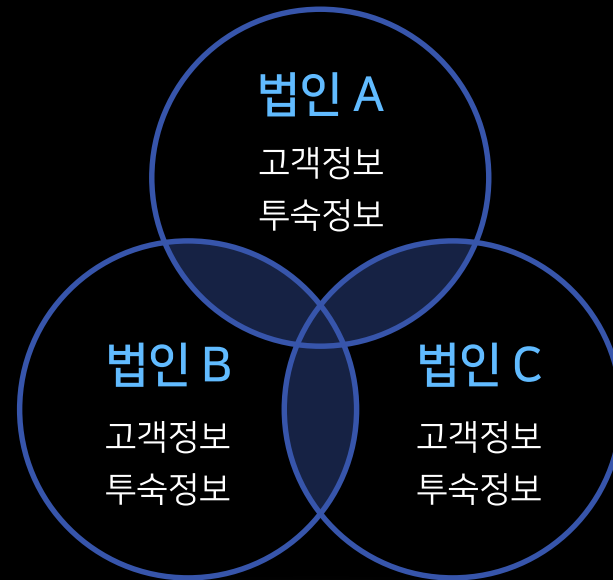
멤버십 도입 전에 가명데이터 결합으로 효과를 예측하고 최종 판단하시죠



Solution

계열사간 고객 정보 결합

고객 통합 및 멤버십 도입 효과 예측



동일 고객 현황 분석을 통해 멤버십 회원의 각 계열사 이용 현황 파악

→ 통합 멤버십 도입 효과 검증

신규 가입 0% ↑

X

회원 비중 확대 00 → 00%

X

교차 이용 증가 00 → 00%

매출 증가

사례 | 제조 유통

대리점 매출 강화를 위해, 고객 거주지 정보와 상권 정보를 분석하여 매출 상승 가능한 지점을 도출해 보았습니다.

Challenge

✓ 고객의 거주 아파트와 구매 제품의 상관관계 확인

✓ 대리점 매출 확대 필요

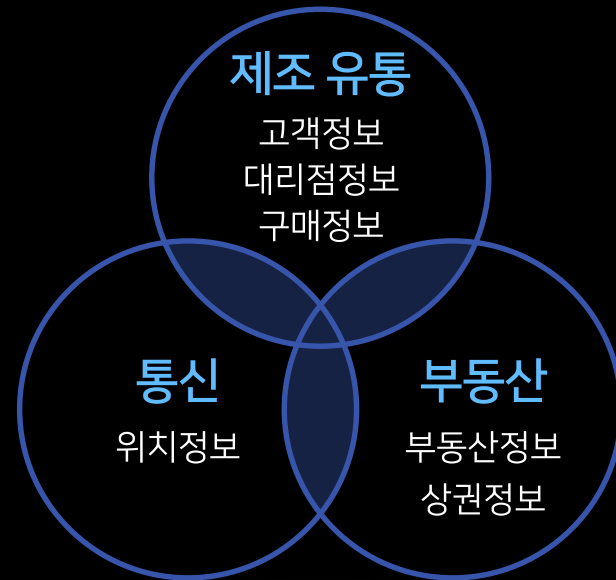
“매장의 위치와 매출이 연관이 있는가?”
“잠재 수요가 높은 지역은 어디인가?”



Solution

통신사 및 부동산 IT사 데이터와 결합

구매 고객의 생활 패턴에 대한 이해



부동산 유형과 Life Style에 따라서 다양한 제품 구매 특성 도출

→ 고객 세분화 강화 및 마케팅 활용

지역별 부동산 유형에 따른 매출 예측

실매출
0,000세대,
000억

예측
모델

추가 매출 예측
0,000세대,
00억

→ 대리점별 성과평가 및 영업관리 활용

예시 | 테마파크

고객의 위치정보로 고객취향을 반영한 제휴마케팅 고도화, 비교 대상 업체 방문 현황 파악이 가능합니다.

Challenge

✓ 고객 선호 기반 **이벤트 기획 및 입점 매장 재배치**

“다른 테마파크 업체의 어떤 것에 우리 방문객들이 관심이 있었을까?”

“우리 고객들은 평상시에 어떤 매장/브랜드를 많이 찾고 있나?”

“우리 고객들이 자주 방문하는 App은 어떤 것이 있을까?”

Solution

통신사 및 유통 멤버십과의 결합

목적에 따라 여러 파트너사와 각각 결합

파악된 고객 특성 기반 서비스 기획



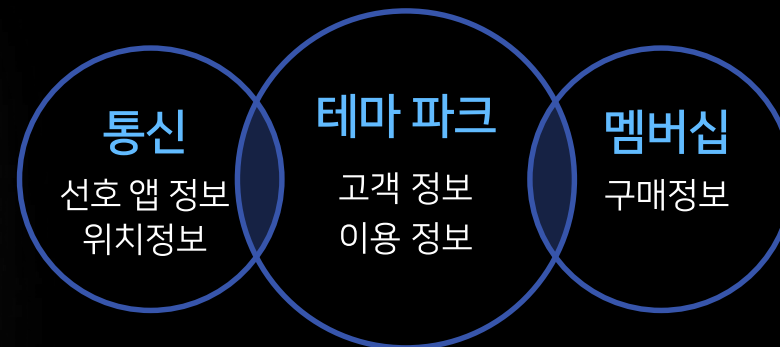
다른 테마파크 이벤트 기간 방문 빈도 고려, 프로모션 기획



자주 가는 상점, 브랜드 파악하여 입점 검토



자주 사용하는 App에 광고나 제휴마케팅 타당성 검토



예시 | 전자 제조

전자제품 구독 고객에 대한 깊은 이해를 통해 구독 고객 확대가 가능하며, 신상품 및 서비스 기획에 도움이 됩니다.

Challenge

- ✓ 구독 관련 신규 고객 확대 및 서비스 차별화 필요

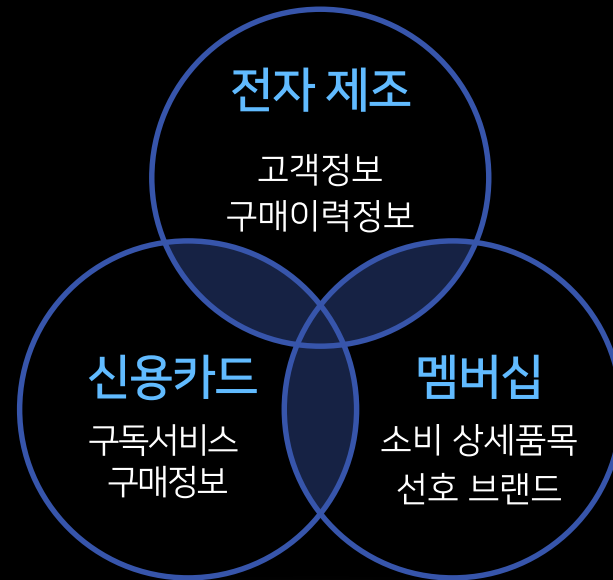
“구독 서비스 고객은 제품만 구매하는 고객과 어떻게 다른가?”

“내가 갖고 있는 고객정보만으로는 고객을 이해하는 데 부족하다”
(이름/성별/주소/구매이력)

Solution

전자제품 구독 고객과 카드/멤버십 결합

기존 고객 이해 및 가망 고객군 파악



구독 고객의 행동/관심 등 특성 이해
→ **인사이트 확보**



구독 가능성이 높은 고객군 도출
→ **마케팅 및 판매 확대**



구독 신상품 및 서비스 기획
→ **차별화·경쟁력 강화**

III.

방법 : 어떻게 결합해야 하는가

결합은 누가하나요? 지정된 전문기관을 통해서만 가능

결합전문기관 (非금융데이터)

지정부처

개인정보위
보건복지부
과기정통부
국토교통부
교육부
산업통상자원부
행정안전부

22개

삼성SDS, LG CNS, SK(주)
통계청, 건강보험심사평가원,
한국지능정보사회진흥원(NIA) 등

데이터전문기관 (금융데이터)

지정부처

금융위원회

4개

금융보안원, 한국신용정보원,
금융결제원, 국세청

2020년 하반기~

민간 1호 결합전문기관으로 지정

① 글로벌 TOP Tier

글로벌 IT 서비스 브랜드 가치 10위
국내 제조 IT 1위

② 안전한 결합

세계최고 암호화 기술 보유
동형암호, 보안경진대회 1위

③ 차별적 가치 창출

Gartner, Forrester가 인정한
글로벌 분석 플랫폼 및
데이터분석가 1,700명 보유

금융데이터 결합 수요 증가로 데이터전문기관 추가 지정 추진 중
민간기업 참여 기회 제공에 따라, **삼성SDS도 신청 참여**

결합 절차는 어떻게 되나요? 1) 사전 준비

1) 사전 준비

2) 데이터 소싱

3) 데이터 결합

4) 데이터 분석

데이터 결합 이해

- ✓ 가명처리 및 결합 제도 이해
- ✓ 데이터 결합 절차 교육

사전 계획 수립

- ✓ 결합 목적 및 활용방안 정의
- ✓ 목적에 맞는 결합 데이터 선정

가명 처리 /
적정성 심사

- ✓ 가명 처리를 위한 기술적 지원
- ✓ 적정성 심사 가이드 및 심사위원 소싱 지원

결합 신청 준비

- ✓ 결합 신청 방법 가이드
- ✓ 결합 데이터 선정, 전처리 등 추출 지원

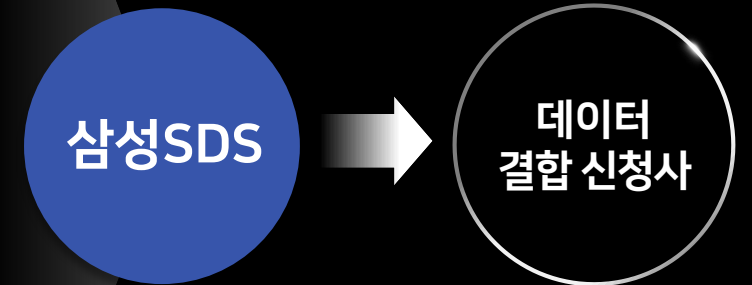
결합 절차는 어떻게 되나요? 2) 데이터 소싱

1 사전 준비

2 데이터 소싱

3 데이터 결합

4 데이터 분석



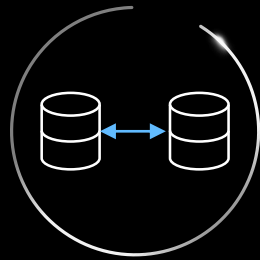
결합 절차는 어떻게 되나요? 3) 데이터 결합

1 사전 준비

2 데이터 소싱

3 데이터 결합

4 데이터 분석



데이터 결합		추가 가명 처리		반출 심사 및 반출	
다양한 결합 방식		다양한 가명처리 방식		반출심사위원회 구성	
다중데이터 결합		KLT*		법률 전문가	
결합률 확인	모의결합	Aggregation	데이터 삭제	보안 전문가	
가명정보 추출	반복 결합	범주화	마스킹	데이터 전문가	

* KLT : k-익명성, l-다양성, t-근접성

결합 절차는 어떻게 되나요? 4) 데이터 분석

1 사전 준비

2 데이터 소싱

3 데이터 결합

4 데이터 분석

✓ 반출된 결합데이터 분석

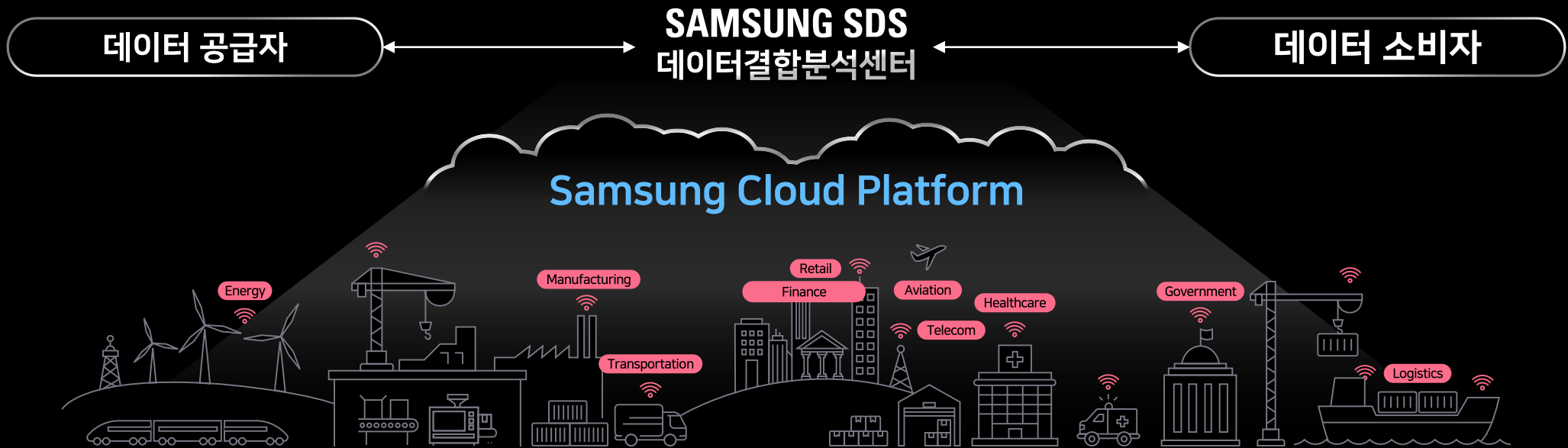
- 삼성SDS 전문 Data Scientist 활용
- 보유 고객군에 대해 결합된 데이터의 상관도 분석 및 예측 모델 개발



✓ Brightics AI 와 클라우드 기반 원격 분석 환경 제공 (예정)



마무리하며... Why Samsung SDS



특장점

- ① 다양한 업종 데이터 경험
- ② 데이터 전문가 (1,700명)
- ③ 데이터분석 플랫폼 (B.AI)
- ④ 삼성만의 강력한 보안

삼성그룹사 및 대외
데이터 컨설팅 경험

데이터 엔지니어 및 데이터
사이언티스트

Gartner, Forrester 등재
(국내 최초)

동형암호, Blockchain,
Endpoint Security 등

Q&A

무엇이든 물어보세요

삼성SDS 데이터결합분석센터 전문가가 도와드립니다.

이메일 : dic.samsung@samsung.com

Thank you

SAMSUNG SDS