


2021 가명정보 활용 우수사례집

 개인정보보호위원회

 과학기술정보통신부

 금융위원회

 보건복지부

NIA 한국지능정보사회진흥원

2021 가명정보 활용 우수사례집



2021 가명정보 활용 우수사례집



개인정보보호위원회



과학기술정보통신부



금융위원회



보건복지부

NIA 한국지능정보사회진흥원

발간사



저희 개인정보보호위원회는 출범과 함께 가명정보 제도를 도입하여 다양한 이종 산업간 데이터 결합을 통해 우리의 정보를 안전하게 활용할 수 있는 기반을 마련하였습니다.

그러나 아직 제도에 대해 잘 알지 못하는 국민들이 많고 현장의 데이터 결합 시도도 활발하지 않은 상황입니다. 다양한 가명정보 결합사례를 발굴하여 관심 있는 기업·기관과 국민들에게 제도를 널리 알리는 것이 중요합니다.

이에 개인정보보호위원회는 가명정보 제도의 효용을 알리고 현장의 활용사례가 늘어나는 데 도움이 될 수 있도록 가명정보 활용 우수사례집을 발간하게 되었습니다.

이번 우수사례집에는 국민생활과 밀접한 선도사례 발굴을 위해 지난해 추진하였던 가명정보 결합 5대 분야 시범사례, 지난 11월에 개최한 '제1회 가명정보 활용 아이디어·우수사례 경진대회'의 수상작품 등 총 17건의 가명정보 결합 우수사례를 담았습니다.

암 환자의 임상정보와 진료정보를 결합하여 암 환자의 장기합병증 및 만성질환을 예측한 사례, 약 340만명 대상의 행정데이터를 결합해 노후소득보장 현황을 분석한 사례 등 본 사례집에 수록된 다양한 사례들은 개인정보 보호의 토대 위에서 데이터를 안전하게 활용할 수 있는 가명정보 제도의 편익을 더 많은 국민들이 체감할 수 있도록 해줄 것입니다. 이러한 측면에서 본 사례집이 데이터를 보다 안전하고 가치있게 활용하고자 하는 모든 이들에게 도움이 되길 기대합니다.

앞으로도 저희 개인정보보호위원회는 가명정보 결합 우수사례를 더욱 폭넓게 발굴하여 가명정보 제도를 활성화시킬 수 있도록 노력을 지속해 나가겠습니다.

끝으로 데이터 부가가치 창출을 위해 가명정보의 제공과 결합에 꾸준히 협력해주시는 관계 부처와 각 기관에 깊은 감사를 드립니다.

2022년 2월

개인정보보호위원회 위원장 **윤 종 인**

소개글

2021 가명정보 활용 우수사례집

경진대회 수상작

안전한 가명정보 활용의 저변을 넓히기 위해 2021년 11월 16일 처음으로 개최된 「2021 가명정보 활용 아이디어·우수사례 경진대회」에 접수된 작품 중, 가명정보 활용사례 부문(대상 4, 우수상 5), 활용 아이디어 부문(대상 1, 우수상 3) 수상작을 수록하였습니다. 관련 영상은 NIA 유튜브 채널(<https://youtu.be/Wo9yEwOTzug>)에서 확인 가능합니다.

- ▣ 주최 : 개인정보보호위원회, 과학기술정보통신부, 금융위원회, 보건복지부
- ▣ 주관 : 한국지능정보사회진흥원

1기 시범사례

가명정보 제도의 효용을 널리 알리기 위해 개인정보보호위원회가 관계부처와 합동으로 발굴하여 추진한 첫 가명정보 결합 시범사례로, 2021년 5월부터 12월까지 총 29개 기관, 31종 데이터를 결합하여 도출한 가명정보 결합 선도사례를 수록하였습니다.

CONTENTS

| 가명정보 결합 우수사례 |

- 복지 노후소득보장 종합연구 보건복지부 사회보장위원회 사무국 **08**
경진대회 수상작 1기시범사례
- 의료·건강 암환자의 합병증 및 만성질환 예측 연구 국립암센터 **10**
경진대회 수상작 1기시범사례
- 의료·건강 폐암환자의 사망동향 및 사망예측 연구 국립암센터 **12**
1기시범사례
- 의료·건강 맞춤형 산림치유 프로그램 분석 국립산림과학원 **14**
1기시범사례
- 통신 불법스팸 실태연구 한국인터넷진흥원 **16**
1기시범사례
- 유통·소비 지역별, 상권별, 상품별 소비패턴 분석 (주)케이티, 롯데멤버스 **18**
1기시범사례
- 유통·소비 편의점 시 상품추천모델 구축 (주)이마트24 **20**
경진대회 수상작
- 금융 수요자별 맞춤형 금융지원정책 수립 (주)국민행복기금 **22**
경진대회 수상작
- 금융 중신용/신평일 고객 신용평가모형 개발 카카오뱅크 **24**
경진대회 수상작

- 금융 코로나 전후 소비·투자 관점의 데이터 결합 및 분석 (주)미래에셋증권, (주)롯데카드 **26**
경진대회 수상작
- 금융 금융소외계층을 위한 머신러닝 모형 개발 하나은행 **28**
경진대회 수상작
- 통신·금융 통신정보를 활용한 사회적 취약층 대상 중금리 대출 지원 (주)케이티 **30**
경진대회 수상작
- 교통 화물차 개인사업자 신용도 평가모델 연구 한국도로공사 **32**
경진대회 수상작

| 가명정보 결합 아이디어 |

- 유통·소비 네트워크 회복을 통한 지역경제 활성화와 청년창업 지원 RE:DATA project **36**
경진대회 수상작
- 복지 교통약자의 이동편의를 위한 대중교통 시설 점검 아이디어 도출 T-BTS **38**
경진대회 수상작
- 복지 보건의료데이터 통합과 직업력 결합을 통한 국민직업건강정보체계 구축 최강이 이김 **40**
경진대회 수상작
- 교육 선택과목별 학생 성취도 대조분석을 통한 효율적인 선택과목 운영 아이디어 도출 최형빈 **42**
경진대회 수상작

가명정보 결합
우수사례

노후소득보장 종합연구

보건복지부 사회보장위원회 사무국

"국민들이 노후대비를 어떻게 하고 있는지 종합적으로 파악하고 싶어요"

고령화 사회가 가속화되고 있는데 국민들은 노후대비를 어떻게 하고 있는 걸까? 사회보장위원회 직원 J씨는 고민에 빠졌다. 다양한 연금 등 노후소득보장 제도가 있지만, 공·사연금 등의 관리주체가 다양하고 기초자료가 분산되어 있어 노후소득보장제도 간 보완관계, 저소득층 노인의 사각지대 분석 등의 면밀한 분석은 다소 어렵기 때문이다.

흩어진 노후소득보장 관련 행정데이터를 결합하면 개인의 가입·수급 현황 분석과 사각지대 등을 보다 면밀히 파악할 수 있고, 데이터에 의한 보다 과학적인 정책분석과 개선방안 제시가 가능할 것 같은데 방법이 없을까?



1. 추진배경 및 필요성

- 노인층의 소득불안정과 빈곤이 우리나라 사회보장제도의 가장 큰 문제로 대두됨
- 현재 공·사연금 등의 관리주체가 상이하고 기초자료가 분산되어 있어 노후소득보장체계의 종합적인 분석이 곤란함
- 노후소득보장 관련 각종 행정데이터를 가명정보로 결합하여 사회보장제도 가입 및 수급 현황을 종합적으로 분석하고 사각지대를 파악하여 데이터 기반의 정책 개선방안을 제시하고자 함

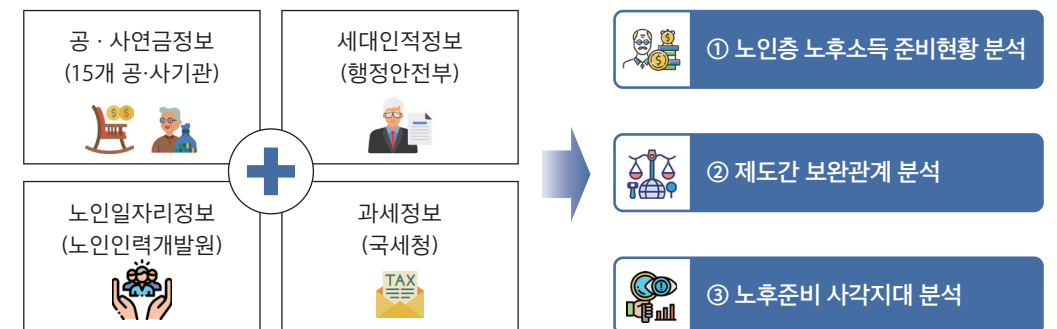
2. 결합데이터

분류	데이터 보유기관	데이터 개요
공공	행정안전부	세대·인적 정보
	국세청	과세정보
	국민연금공단	국민연금 정보
	한국사회보장정보원	기초연금 정보
	공무원연금공단	공무원연금 정보
	국군재정관리단	군인연금 정보
	국방전산정보원	
	별정우체국연금관리단	별정우체국연금 정보
	사립학교교직원연금공단	사립학교교직원연금 정보
	한국주택금융공사	주택연금 정보
	한국농어촌공사	농지연금 정보
	한국노인인력개발원	노인일자리 정보
민간	국민·기업·농협·신한·우리·하나 은행	퇴직연금 정보

▶ 결합기관 : 국세청

3. 분석결과 및 향후 계획

- 65세 이상 노인층의 연금수급률은 47%, 연평균 수급액 710만원으로 분석됨
- 65세 이상 남성의 연금수급률 및 평균 연간수급액이 여성 대비 약 2배로 나타남
- 연령이 높아질수록, 연금수급액이 낮을수록 기초연금 수급률이 증가함
- 해당 분석결과를 기반으로 노후소득보장제도 간 보완관계를 분석하고, 저소득층 노인의 소득보장실태 파악 및 사각지대 분석 등 보다 심층적인 분석을 수행할 예정임



암환자의 합병증 및 만성질환 예측 연구

국립암센터

"암 경험자가 특별히 조심해야 하는 합병증이 있나요?"

30대 초반의 M씨는 최근 위암 진단을 받아 수술을 마쳤다. 다행히 초기에 발견하여 심각한 상황은 면했지만 앞으로의 건강관리에 대한 걱정이 많아졌다. 암 경험자들이 모인 인터넷 카페에서 여러 정보를 검색해보아도 파편화된 정보만으로는 일반화가 가능한 정보인지 단순 개인의 사례인지 판단하기가 어렵다.

M씨는 위암 경험자에게 특히 자주 발생하는 합병증은 없는지, 특별히 조심해야 하는 만성질환은 없는지 미리 알고 그에 대비하여 건강관리를 하고 싶다.



1. 추진배경 및 필요성

- 최근 조기진단 및 치료기술 발달에 따라 암 생존자의 치료 이후 발생하는 장기적인 합병증과 만성질환 관리를 통한 '치료 후 관리'의 중요성이 대두됨
- 암 치료가 끝난 환자가 다른 병원을 이용할 경우, 단일병원 임상정보만으로는 환자의 장기 합병증, 만성질환 발생 여부 등을 확인할 수 없음
- 암환자의 임상자료, 건강보험자료를 결합하면 암 치료효과를 분석하고 암환자에게 주로 발생하는 장기 합병증과 만성질환, 사망에 관련된 유의미한 결과를 도출할 수 있음
- 암 생존자에게 발생할 수 있는 부작용을 사전에 예측하고 최적의 치료법 제공이 가능해짐
- 암 생존자의 만성질환 관리 및 임상의료 효율 증대를 통해 보건의료 비용도 절감할 수 있음

2. 결합데이터

데이터 보유기관	데이터 개요	레코드 수
국립암센터	암환자 및 일반환자의 임상정보	약 20만명
국민건강보험공단	암환자 및 일반환자의 장기추적 관찰 진료정보	약 20만명

▶ 결합기관 : 국민건강보험공단

3. 분석결과 및 향후 계획

- 6대 암환자가 일반 환자군에 비해 심뇌혈관질환 중 심부전(81% ↑), 심근경색(50% ↑), 뇌졸중(25% ↑) 발생 빈도가 더 높았으며, 특히 심부전의 발생 빈도가 심뇌혈관질환 중에서 가장 높게 나타남
- 근골격계질환 중에서 골절은 6대 암 환자에서 일반 환자군에 비해 47% 발생이 더 많았고, 대사질환 중에서 가장 대표적인 당뇨병의 발생은 35% 더 많았음
- 이외 간암·갑상선암 환자의 경우 일반 환자군에 비해 심뇌혈관질환 및 근골격질환 발생 위험도에 큰 차이가 없었음
- 향후 심층분석을 통해 6대 암종별 장기 합병증과 만성질환의 세부 발생현황 및 주요 발생요인을 파악하고, 결합데이터의 AI학습을 통해 암 생존자의 생애주기 전반에 걸친 장기질환에 대한 위험요인을 파악하고 예방을 위한 예측 모델을 제시할 계획임



폐암환자의 사망동향 및 사망예측 연구

국립암센터

"폐암 생존자가 가장 조심해야하는 만성질환은 뭐예요?"

직장인 C씨는 최근 폐암 수술을 받고 항암치료를 마친 단짝친구를 위해 의미있는 선물을 하고 싶다. 아직 젊은 나이이니 앞으로 특히 조심해야 하는 만성질환이 있다면 그에 대비할 수 있도록 좋은 영양제를 선물하기로 결정했다.

그런데 폐암 생존자가 어떤 질환에 가장 취약한지, 치료 예후는 평균적으로 어떤지 파악하고 그에 맞는 영양제를 구입하고 싶어 인터넷을 검색해보니 너무 많은 정보들이 쏟아진다. 어떤 게 믿을만한 정보인지 판별하는 것도 어렵다.



1. 추진배경 및 필요성

- 암 환자는 해당 암 외 다양한 질환의 발생 및 사망위험에 노출되므로, 생애주기 전반에 걸친 위험요인 파악 및 맞춤형 진단·치료로 장기 생존율을 제고할 필요가 있음
- 환자들이 여러 병원을 이용할 가능성을 배제할 수 없어 단일 병원 데이터만으로는 추적관찰 등에서 주요 결과를 파악하는 데 한계가 있음
- 여러 기관이 보유하고 있는 건강 관련 빅데이터를 가명처리 후 결합하여 폐암 치료효과를 분석하고 폐암 환자에서의 합병증, 만성질환 발생 및 사망 예측모델을 개발하고자 함

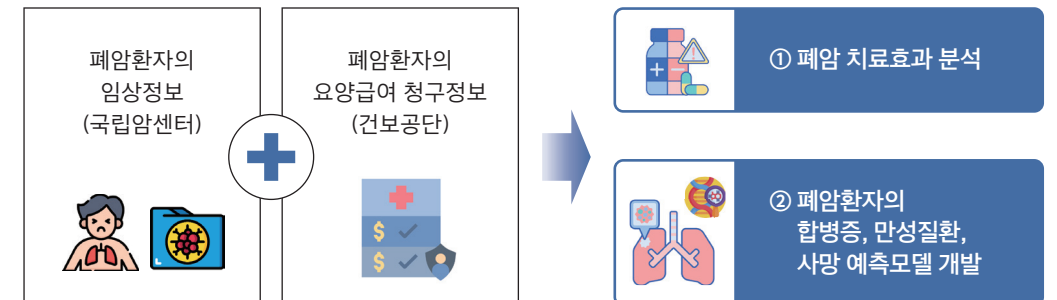
2. 결합데이터

데이터 보유기관	데이터 개요	레코드 수
국립암센터	폐암 환자의 진단, 검사, 병리, 수술, 항암제 등 임상 정보('02년 ~ '19년)	약 2만명
국민건강보험공단	폐암 환자의 건강검진, 상병내역, 요양기관 등 요양 급여 청구 정보('11년 ~ '19년)	약 2만명
통계청	사망연월일 및 사망원인 103항목 분류 정보('04년 ~ '19년)	약 423만명

▶ 결합기관: 통계청

3. 분석결과 및 향후 계획

- 국립암센터에 내원한 폐암환자 1만4천여 명 중 1년 이내 사망은 38.2%, 3년 이내 사망은 67.3%, 5년 이내 사망은 77.4%, 10년 이내 사망은 87.5%로 분석됨
- 폐암 진단을 받고 5년 이상 생존 후 연구대상기간 내 사망한 환자의 22.2%가 암 이외의 원인으로 사망하였고, 이 중 심뇌혈관질환으로 인한 사망이 가장 많은 24.8%를 차지함
- 향후 폐암 환자의 단기·중기·장기 사망원인 및 연도별 사망동향을 파악하고, 심층분석을 통해 폐암 환자에서의 만성질환 발생 및 사망과의 인과관계를 확인할 예정임



맞춤형 산림치유 프로그램 분석

국립산림과학원

"건강관리를 위해 효과가 좋은 운동을 알고 싶어요"

30대 직장인 L씨는 최근 받은 건강검진 결과에 큰 충격을 받았다. 더 이상은 젊음을 믿고 건강관리에 소홀해서는 안되겠다는 생각에 적당한 운동을 통해 체중을 감량하고 내년 검진에서는 모든 검사수치를 정상범위 내로 돌려놓는 것을 목표로 잡았다.

막상 운동을 시작하려니 어떤 운동을 선택해야 좋을지 고민되기 시작했다. 내 또래의 남들은 보통 어떤 운동을 하고 있는지, 어떤 운동이 가장 적은 노력으로 큰 체중감량 효과가 있는지 알고 싶다.



1. 추진배경 및 필요성

- 산림치유 효과성 연구는 소규모 인원을 대상으로 하고 단기 효과검증 연구에 집중되어 있어 장기적인 산림치유 효과의 의학적 유효성 검증에 한계가 있음
- 산림치유 효과에 대한 의과학적 근거를 마련하고 숲이 주는 건강효과의 정량적 평가가 가능하다면 산림 의료연계 서비스 개발에 활용할 수 있을 것으로 기대됨
- 인체 건강영향 검증과 보건의료서비스를 결합·분석하여 보다 나은 산림치유 서비스를 제공하고 이를 통해 국민 의료비 절감에 기여하고자 함

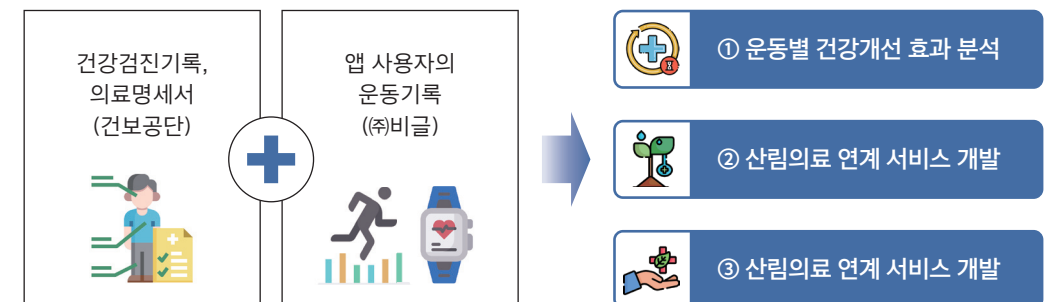
2. 결합데이터

데이터 보유기관	데이터 개요
국민건강보험공단	일반건강검진기록 및 의료명세서 정보('11~'20년)
(주)비글	앱 사용자의 운동기록('11~'20년) - 활동에 따른 누적거리(등산, 걷기, 뛰기, 자전거), 활동 등록수(등산, 걷기, 런닝, 기타), 운동종목코드, 이동거리, 운동거리, 평균속도, 운동시간 등

▶ 결합기관 : 국민건강보험공단

3. 분석결과 및 향후 계획

- 연령이 높아질수록 운동(등산·걷기·달리기·자전거) 중 등산을 많이 하고 달리기의 비율은 감소함
- 모든 운동에서 운동량과 빈도가 증가할수록 혈압, 콜레스테롤, 공복혈당 등 대부분의 건강지표가 정상 범위로 안정화되었고, 입내원, 요양일수, 총의료비 등 의료부담도 감소하는 경향을 보임
- 특히, 체중·체질량·허리둘레 등의 건강지표의 개선에는 다른 운동보다 등산의 횟수나 거리가 큰 영향을 주는 것으로 나타남
- 분석결과를 기반으로 녹지율, 산림면적 등 산림 빅데이터 연구를 통해 숲에서의 운동이 건강에 미치는 영향을 심층분석할 예정임



불법스팸 실태연구

한국인터넷진흥원

"불법스팸문자, 저만 이렇게 많이 오는 건가요?"

직장인 H씨는 요즘 스팸과의 사투를 벌이는 중이다. 온갖 대출권유, 주식투자정보, 도박권유 등 전화며 문자며 하루에도 몇 번의 스팸을 받고 있어 스트레스가 이만저만이 아니다. 휴대전화번호를 바꿀까도 생각해보았지만 직업 특성상 이 또한 쉽지 않다.

H씨는 본인만 유독 불법스팸으로 괴로움을 겪고 있는 건지, 남들도 나와 비슷하게 다양한 유형의 스팸을 받고 있는 건지 불법스팸의 전반적인 실태가 궁금해졌다.



1. 추진배경 및 필요성

- 수신자의 동의 없이 일반적으로 전송되는 불법스팸은 많은 국민들에게 정신적·금전적 피해를 주고 있어 대응정책 수립이 절실함
- 한국인터넷진흥원 불법스팸대응센터로 신고되는 불법스팸정보에는 광고번호, 내용, 유형 등 단편적인 정보만 포함되어 수신자의 연령, 성별 등에 따른 불법스팸 실태를 파악하기 어려움
- 한국인터넷진흥원의 스팸신고정보와 통신사의 가입자 정보를 결합하여 성별, 연령대별 불법스팸 피해 실태를 파악하고 연령대별 맞춤형 스팸예방 교육과 인식제고 사업 추진의 과학적 기반을 마련하고자 함

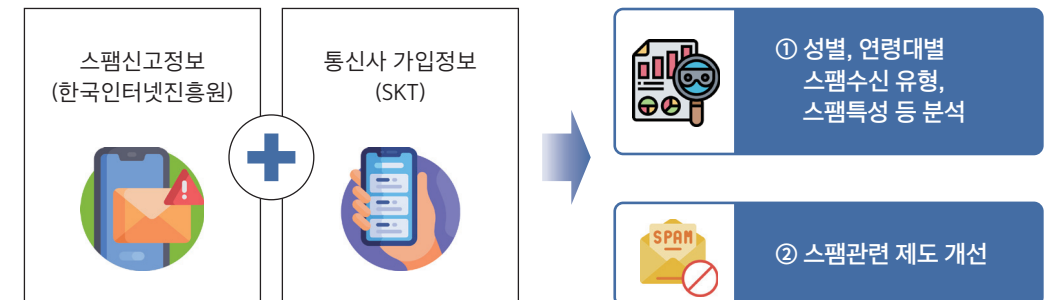
2. 결합데이터

데이터 보유기관	데이터 개요
한국인터넷진흥원	스팸신고정보 - 발신번호, 수신일시, 신고일시, 신고건수, 발송유형(문자, 음성), 분야(도박, 대출 등)
SKT	통신사 가입정보(고객 연령, 성별 등)

▶ 결합기관: 삼성SDS

3. 분석결과 및 향후 계획

- 스팸신고 비율은 남성 64.4%, 여성 35.6%로 남성이 여성에 비해 약 1.8배 높았으며, 연령대별로는 50대 28.6%, 60대 22.8%, 40대 22.7% 순으로 높은 비율을 차지함
- 대부분의 성별·연령대에서 문자스팸 비율이 음성스팸에 비해 높았으나, 20대 미만·20대·40대 여성의 경우 음성스팸 비율이 근소하게 높은 것으로 나타남
- 스팸유형 중 도박스팸이 차지하는 비율이 가장 높았고 이어 불법대출 순이었으며, 남성은 여성에 비해 주식정보 스팸, 여성은 남성에 비해 대출·카드스팸 비율이 상대적으로 높았음
- 결합 데이터 분석 결과를 스팸탐지기술 정교화 등 스팸관련 제도·시스템 개선에 활용할 예정임



지역별, 상권별, 상품별 소비패턴 분석

(주)케이티, 롯데멤버스

"아이 키우는 집은 마트에서 주로 어떤 걸 구입하나요?"

주부 J씨는 마트에서 장을 볼 때 다른 사람들의 카트 안에 어떤 물건이 실려있는지 구경하는 걸 좋아한다. 여러 카트에 공통적으로 실려있는 물건은 어느정도 품질이 보증된 물건 같아 J씨도 슬그머니 카트 안에 하나 넣어보곤 한다. 카트 안의 물건을 보면 '저 사람은 아이를 키우는 아빠겠구나', '저 사람은 혼자 사는 사람인가 보다' 대강 추측해 보는 재미도 있다.

실제로 가족 구성원이나 기타 특징에 따라 소비행태가 모두 다를 텐데 이처럼 고객 특성에 따른 소비행태 분석이 있다면 재미있을 것 같다는 생각이 들었다. 하지만 마트에서는 고객의 가구속성이나 거주지 등의 정보를 갖고 있지 않을 테니 그런 분석은 어려우려나?



1. 추진배경 및 필요성

- 소비자 개인별 · 지역별 특성에 따른 다양한 수요에 대응한 맞춤형 상품과 서비스를 제공한다면 소비자와 기업 모두 얻을 수 있는 편익이 크지만, 각 기업에서 보유한 데이터 분석만으로는 고객 형태별 소비행태 분석에 한계가 있음
- 유통사의 구매정보와 고객 개인 특성 정보를 결합하면 세대 구성별 · 지역별 · 구매금액별 소비패턴을 파악하여 그에 맞는 맞춤형 상품과 서비스 판매전략을 구상하고 이용자들이 원하는 상품을 추천할 수 있음

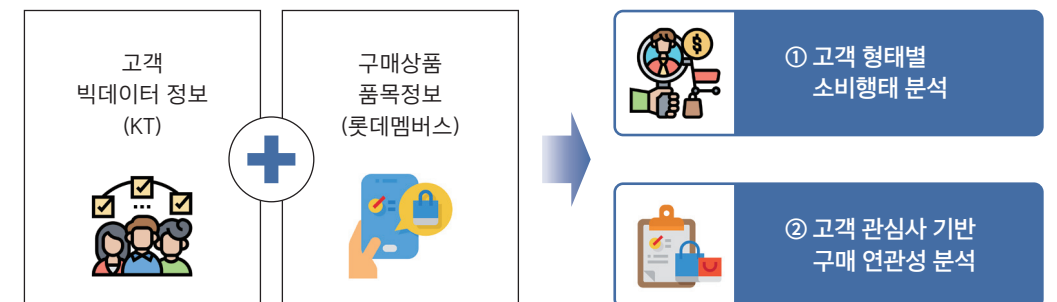
2. 결합데이터

데이터 보유기관	데이터 개요
(주)케이티	고객 성별 · 연령별 정보, 빅데이터 분석을 통해 도출한 거주지, 근무지, 관심분야, 가구구성 형태 등 추정정보
롯데멤버스	서울지역 롯데마트 · 하이마트 구매상품의 품목정보, 구매금액 정보

▶ 결합기관 : 한국지능정보사회진흥원

3. 분석결과 및 향후 계획

- 자녀가 있는 가구는 주로 식품을 구매하고, 1인가구는 의류 · 패션잡화, 가전 · 디지털기기의 구매 비중이 높았음
- 구매금액 기준 상위 고객은 30대이며, 유 · 아동 의류와 완구의 구매 비중이 높은 것으로 확인됨
- 개인별 관심도가 높은 품목에 대한 구매 비중이 2배 이상 높게 나타남



편의점 AI 상품추천모델 구축

(주)이마트24

"편의점도 소비자 맞춤형 판매전략을 수립하고 싶어요"

편의점을 운영하는 점주 A씨는 최근 방문객과 매출이 급격하게 줄어들어 근심이 많다. 수년간의 운영경험을 바탕으로 상품구성과 진열도 달리해 보았지만 이렇다 할 효과를 보지 못했다. 고객들의 소비패턴과 취향을 파악할 수 있다면 원하는 상품을 맞춤형으로 진열하고 상권에 맞는 재고관리 전략을 수립할 수 있을 텐데 오프라인 매장 특성상 어려워 보인다.

점주들의 깊어지는 고민을 해결하기 위해 프랜차이즈 B사는 편의점이 보유한 점포별 영수증 데이터와 카드회사가 보유한 고객의 소비패턴 정보를 결합하는 방법을 떠올렸다. 상세한 고객 소비 행태 분석을 통해 개별점포의 상권, 주 방문고객 선호도 등을 고려한 가맹점 맞춤형 컨설팅이 가능하지 않을까?



1. 추진배경 및 필요성

- 오프라인 유통, 특히 편의점은 지점을 방문한 고객에 대한 데이터 확보가 어려움
- 이를 보완하고자 편의점 업계에서도 자체 모바일앱을 통해 고객 정보를 활용하려 하지만 결제 시 포인트를 적립하는 경우에만 고객정보를 얻을 수 있어 한계가 있음
- 고객 결제정보를 보유한 카드사의 데이터와 그 카드를 보유한 고객이 편의점에서 구매한 상품 데이터를 결합하여 분석하고, 이를 기반으로 고객 맞춤형 상품을 제공하고자 함

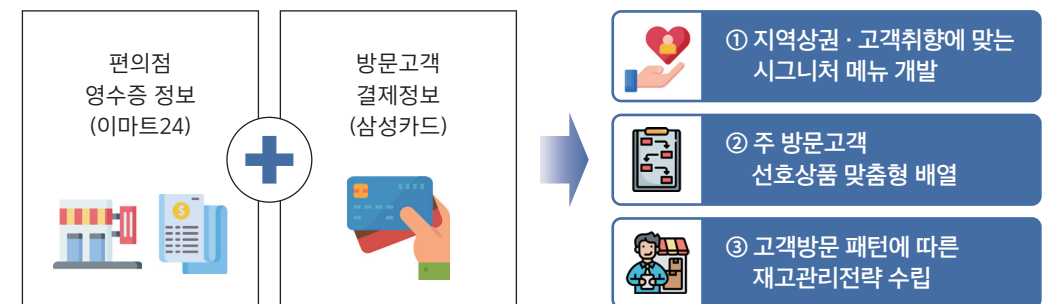
2. 결합데이터

데이터 보유기관	데이터 개요	레코드 수	칼럼 수	데이터크기
(주)이마트24	삼성카드로 결제된 이마트24 거래(영수증) 데이터 : 거래 및 품목 정보	15,789,954개	157개	6.1GB
(주)삼성카드	삼성카드의 이마트24 매출정보 및 이용고객 / 가맹점 특성 정보	16,127,415개	69개	6.0GB

▶ 결합기관 : 금융보안원

3. 분석결과 및 향후 계획

- 고객 맞춤형 상품구성으로 고객 편의성을 높이고 점포의 운영역량 및 브랜드 이미지를 제고할 수 있음
- 고객 만족도 제고로 단골 · 신규고객 확대 및 점포의 가치상승이 기대됨



수요자별 맞춤형 금융지원정책 수립

(주)국민행복기금

"최저신용자를 대상으로 대출상품을 구상중인데 금리를 어떻게 설정해야 하죠?"

K기금에서는 다양한 맞춤형 서민금융상품을 운영중이다. 정부의 법정 최고금리 인하 결정에 따라 기존 운영하던 대출상품의 금리 조정이 필요하게 되었다.

법정 최고금리가 인하되면 서민들의 이자부담이 줄어들어 좋은 것 아닌가 싶지만, 줄어든 이자 수익을 메꾸기 위해 금융사에서는 상대적으로 돈을 갚지 못할 확률이 높은 저신용자의 대출을 대폭 줄이게 되고, 저신용자들은 돈을 빌리지 못해 불법 사금융으로 내몰릴 수 있다.

이를 방지하기 위해 K기금에서는 최저신용층이 이용 가능하면서도 금융시장 왜곡을 최소화 할 수 있는 적정 금리를 정하고, 대출수요를 파악하여 한정된 재원으로 효과적인 지원을 할 수 있는 금융지원 정책을 수립하고자 한다.



1. 추진배경 및 필요성

- 정부의 법정 최고금리 인하('21. 7.7., 24% → 20%)에 따라 서민의 금융이용 축소 등의 부작용을 최소화하기 위하여 정책서민금융 공급확대 등 다각적인 대응방안을 마련하고자 함
- 제도권 금융 이용이 거절되어 불법 사금융 등 고금리대출 이용이 불가피한 이들을 지원하기 위한 정책상품을 설계하고, 수요예측 등을 위해 금융취약계층의 대출이용행태를 분석할 필요성이 대두됨
- 신용집중기관이 보유한 대출금리 및 잔액정보와 CB사가 보유한 신용평점 정보 등을 결합하여 금융 취약계층의 대출이용 행태를 분석하고, 기존 정책상품의 금리체계 개편 및 신규 대출상품 출시를 기획함

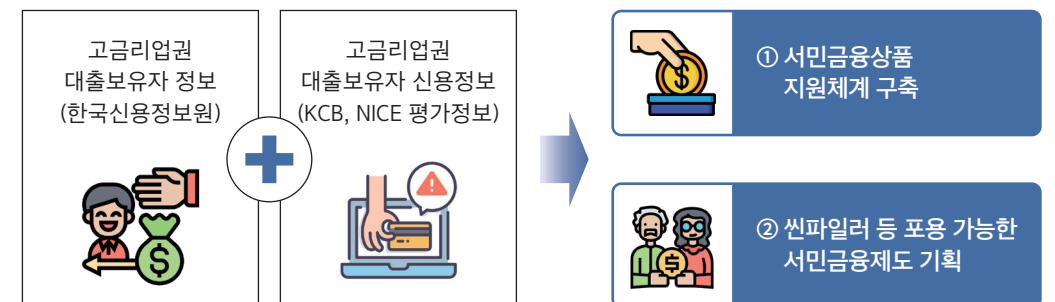
2. 결합데이터

데이터 보유기관	데이터 개요	레코드 수	칼럼 수	데이터크기
한국신용정보원	고금리업권(카드·여신전문·저축은행·대부업권) 대출 보유자의 대출금리, 연령, 성별, 소득, DSR, 연체여부 등의 정보	8,914,429개	10개	0.786 GB
NICE평가정보(주)	고금리업권(카드·여신전문·저축은행·대부업권) 대출 보유자의 신용평점 및 하위누적구성비 등의 정보	3,535,030개	5개	0.277 GB
코리아크레딧뷰로(주)	고금리업권(카드·여신전문·저축은행·대부업권) 대출 보유자의 신용평점 및 하위누적구성비 등의 정보	5,472,277개	5개	0.410 GB

▶ 결합기관 : 한국신용정보원

3. 분석결과 및 향후 계획

- 기 운영중인 정책상품 햇살론17의 금리체계를 개편하고 안전망 대출 II 신규 출시를 기획하는 등 서민·실수요자 위주의 서민금융상품 지원체계를 구축함
- 금융취약계층 뿐만 아니라 금융정보가 부족한 Thin-Filer를 위한 지원체계를 구축하여 보다 다양한 대상을 포용하는 서민금융제도를 기획할 예정임



중신용/신패일 고객 신용평가모형 개발

카카오뱅크

"금융소외계층 고객들에게도 폭넓은 지원을 줄 수 있는 방법이 있을까요?"

K은행 H과장은 중신용, 금융소외계층 고객들에게 대출을 확대할 수 있는 방법을 찾아 나섰다. 신용평가에 활용할 수 있는 비금융정보가 확대되고 있음에도 여전히 금융정보 중심의 평가가 이루어지다 보니, 금융이력부족 고객은 신용도가 높을지라도 대부분 중위등급을 부여 받는 등의 불이익을 받는 경우를 많이 보았기 때문이다.

금융소외계층의 신용도를 정확하게 평가할 수 있도록 비금융정보를 활용한 신용등급 재평가 모델을 개발하면 신용도가 양호한 금융소외계층이 은행 대출에서 불이익을 받지 않게 되고 은행 입장에서도 대출가능 고객군을 확대할 수 있어 모두에게 좋은 개선이 될 것 같다.



1. 추진배경 및 필요성

- 기존의 금융정보 위주 신용평가모형으로는 리스크 차별화가 어려웠던 금융이력부족 고객의 리스크 변별을 위하여 대안정보 활용의 필요성이 높아짐
- 다양한 대안정보와 고객의 내부거래 실적정보를 결합하여 신용평가모형을 정교화하고자 함

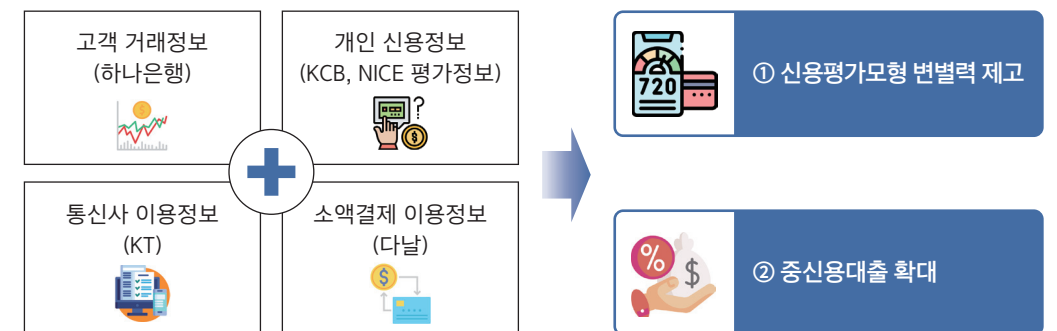
2. 결합데이터

데이터 보유기관	데이터 개요	레코드 수	칼럼 수	데이터크기
하나은행	대출신청 / 개설 고객의 신청 및 거래 데이터	25,865,510개	387개	24GB
코리아크레딧뷰로(주)	카카오뱅크 대출신청 / 개설 고객의 신용거래 데이터	25,791,211개	256개	29GB
NICE평가정보(주)	카카오뱅크 대출신청 / 개설 고객의 신용거래 데이터	25,666,805개	673개	24.9GB
(주)케이티	통신상품 가입 / 이용 및 요금청구 / 수납데이터와 통신정보 기반 신용평가등급	58,172,187개	28개	7.6GB
다날	소액결제 이용자의 거래실적(결제건수 / 금액 / 회차 / 사용영역 등) 데이터 및 다날 소액결제 데이터로 평가한 신용평가등급	9,927,281개	95개	3GB

▶ 결합기관: 금융보안원

3. 분석결과 및 향후 계획

- 개인사업자 매출액정보, 간편결제 및 송금정보, 연말정산 등의 대안정보를 추가적으로 가명결합하여 모형 변별력 개선 및 중신용 대출 확대를 위한 노력을 지속할 예정임
- 모빌리티 기반 종사자인 대리기사, 택시기사의 운행이력정보와 이커머스 셀러의 매출 흐름, 고객리뷰, 반품률 등을 활용하여 모빌리티 종사자 및 소상공인 이커머스 셀러 등 개인사업자에 대한 신용평가 모델도 고도화할 예정임



코로나 전후 소비·투자 관점의 데이터 결합 및 분석

(주)미래에셋증권, (주)롯데카드

"코로나19로 사람들의 소비·투자 패턴이 어떻게 바뀌었을까요?"

직장인 N씨는 매달 조금씩 저축하여 매년 1회씩 해외여행을 다녀오곤 했다. 그러나 코로나19 발생 이후 해외여행이 어려워지자 통장에는 2년 치의 여행자금이 쌓이게 되었다. N씨는 그 돈으로 뭘 할지 고민하다가 올 한해도 고생한 스스로에게 주는 선물로 그간 어마어마한 가격 때문에 눈독만 들이고 있던 명품시계를 하나 구입했다.

N씨처럼 코로나19 전후로 많은 사람들의 소비·투자 패턴이 달라졌을 것이다. 2년이 넘도록 지속된 코로나19 시대에 사람들의 소비·투자 패턴은 어떻게 바뀌었을까?



1. 추진배경 및 필요성

- 코로나19의 영향으로 소비·투자 관련 변화된 양상을 종합적으로 파악하여 국민에게 유용한 정보를 제공하고자 함
- 국민체감성이 높은 카드사의 소비 데이터와 투자사의 고객 투자데이터를 결합·분석하여 국민 생활 편의 개선 아이디어 도출에 기여하고자 함

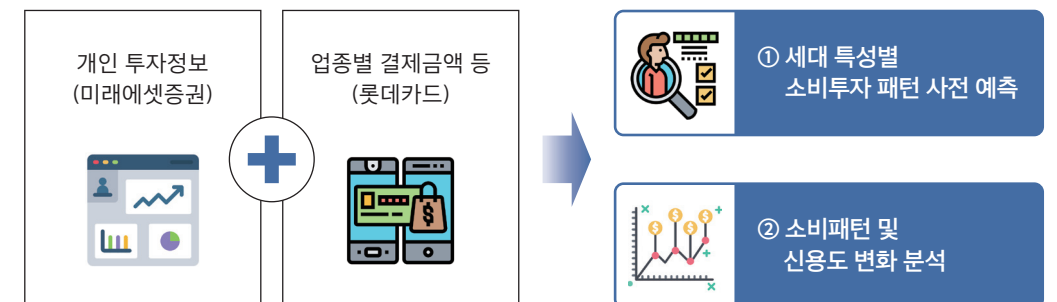
2. 결합데이터

데이터 보유기관	데이터 개요	레코드 수
(주)미래에셋증권	개인 고객의 투자자산 및 약정금액, 주식 거래량 등	약 70만 개
(주)롯데카드	업종별 결제금액, 추정소득, 신용등급 등	약 70만 개

▶ 결합기관: 금융결제원

3. 분석결과 및 향후 계획

- MZ세대는 총 자산 평균이 그 외 세대에 비해 더 많이 증가하였고, 해외주식 비중이 상대적으로 더 큼
- MZ세대는 CMA의 비중이 높으나, 그 외 세대의 경우 주식투자 비중이 더 높음
- 세대별 소비 비중에는 큰 변화가 없으나 소비 금액이 큰 폭으로 감소함
- 비대면 업종의 소비가 증가했고 배달 사용 비중 및 금액이 모두 증가함
- 마트·편의점, 학원 업종의 소비 비중은 비슷했고, 여행·음식·공연·취미 분야의 사용 금액은 큰 폭으로 감소함
- 대출, 투자, 소비 분석을 통해 고객의 수익률, 소비패턴 및 신용도 변화 분석을 진행하고, 데이터 분석 및 거래플랫폼을 통해 누구나 쉽게 연구결과 및 데이터를 활용할 수 있도록 제공할 예정임



금융소외계층을 위한 머신러닝 모형 개발

하나은행

"사회초년생은 신용도가 낮아 대출이 너무 어려워요"

사회초년생 A씨는 집과 직장이 너무 멀어 직장 주변에서 자취를 결심했다. 월세 보증금 마련을 위해 은행에서 대출상담을 받았는데, A씨와 같은 사회초년생은 금융사용 이력이 부족하여 신용도가 낮아 알아보고 온 저금리 대출은 어렵겠다는 답변을 받았다.

A씨는 본인의 상환능력이 양호한데도 신용카드 사용이력, 대출경험이 없다는 이유로 신용도가 낮게 측정된 것이 안타까웠다. 신용평가에 적용할 수 있는 통신정보, 소액결제정보 등 다양한 비금융정보를 신용평가에 반영한다면 금융소외계층*의 신용도도 정확하게 평가할 수 있게 될 텐데..



* 금융소외계층 : 신용평가시점 기준으로 최근 2년간 신용카드 사용이력 및 3년간 대출경험이 없는 금융이력부족자 (Thin-Filer)로, 사회초년생·주부·고령층 등이 해당됨

1. 추진배경 및 필요성

- 사회초년생, 주부, 노년층 등 신용평가정보가 부족하여 정확한 신용평가가 어려운 금융소외계층을 대상으로 포용적 금융을 실현하기 위해 다양한 대안정보를 이용한 Thin-Filer 머신러닝 모형을 개발하고자 함
- 신기술 시 머신러닝 알고리즘을 이용한 정교한 신용평가모형 개발을 위해 다양하고 새로운 대량의 가명정보 결합이 필요함

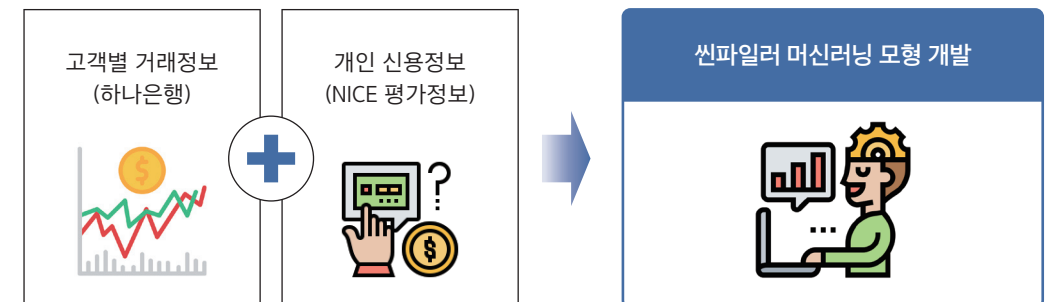
2. 결합데이터

데이터 보유기관	데이터 개요	레코드 수	칼럼 수	데이터크기
하나은행	월별 고객별 수신 및 부수거래 정보	363,859,675개	110개	90GB
NICE평가정보(주)	대출조회 개인에 대한 신용정보	45,181,177개	809개	136GB

▶ 결합기관 : 한국신용정보원

3. 분석결과 및 향후 계획

- 다양한 영역의 대안정보를 추가결합하여 신용평가 머신러닝 모형을 지속적으로 고도화할 예정임
- 신용평가 정교화를 통해 금융소외계층에게는 금융의 문턱을 낮추고 기존 금융소비자에게는 보다 좋은 조건의 대출을 제공하고자 함
- 그룹 내 업권 간 연계대출을 활성화하여 기존 1금융과 2금융간 금리단층 영역인 중신용·중금리 대출을 세밀하게 제공할 예정임



통신정보를 활용한 사회적 취약층 대상 중금리 대출 지원

(주)케이티

"통신비 연체 없는 통신사 장기 가입고객은 신용도도
높을까요?"

직장인 W씨는 사회초년생 시절 금융거래이력이 부족하여 대출을 거절당한 경험이 있다. 사회초년생, 주부 등은 금융거래이력 부족으로 신용도가 낮아 대출이 어렵다는데, 이를 대체하여 신용도 평가에 반영할 수 있는 요소들이 있지 않을까?

예를 들어 휴대폰은 대부분의 사람들이 사용하므로 요금납부 이력, 가입기간 등의 정보가 있을 것이다. 통신비를 연체 없이 납부하고, 잦은 해지 없이 장기 가입되어 있는 고객은 상대적으로 경제적으로 어려움을 겪고 있지 않을 확률이 높을 것 같은데, 이렇게 통신정보를 신용평가에 반영할 수 있지 않을까?



1. 추진배경 및 필요성

- 기존 금융정보 기반의 개인신용평가는 직장경력, 기존 소득 등을 위주로 개인 신용도를 평가하여 사회 초년생, 주부 등은 변제능력이 있음에도 불구하고 대출시장에서 불이익을 받는 문제가 발생함
- 통신정보는 고객의 지불의향과 상환능력을 예측할 수 있는 대안정보로, 통신정보를 결합하면 보다 정교하게 고객을 식별할 수 있어 기존 금융소외계층인 싹파일러(Thin-File)를 위한 대안신용평가 모형 개발이 가능함

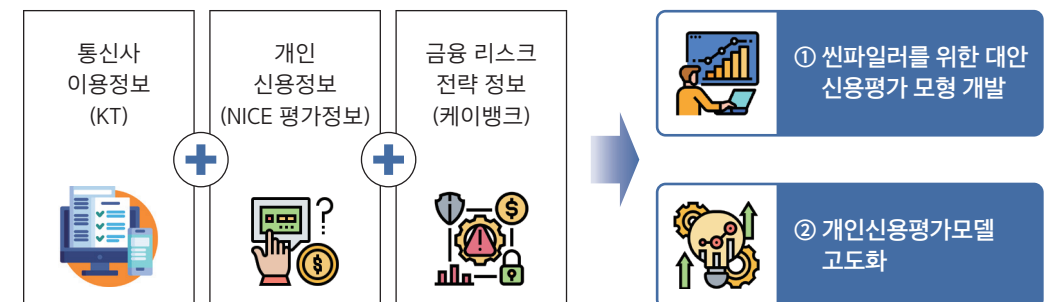
2. 결합데이터

데이터 보유기관	데이터 개요	레코드 수
(주)케이티	KT 통신정보를 조회한 케이뱅크 고객대상 TS 및 TPS 항목 범주화 데이터	약 100만 개
NICE평가정보(주)	케이뱅크 조회 대상자에 대한 신용정보 데이터	약 200만 개
케이뱅크	케이뱅크 승인대상자에 대한 리스크 전략 적용 데이터	약 200만 개

▶ 결합기관: 금융보안원

3. 분석결과 및 향후 계획

- 기존 금융거래이력이 부족하여 신용대출에 어려움이 있었던 고객군 대상으로는 새로운 중금리 대출 지원이 가능하게 되고, 반대로 통신이력이 상대적으로 불량한 고객군은 신용대출 등 금융혜택을 일부 제한하여 국내 대출시장의 효율화 및 건전화에 기여하게 됨
- 보다 정교화된 신용평가모형으로 연체 등 금융사 리스크를 낮춤으로써 금융사의 손실절감 및 우량 고객 대상 추가 금융혜택 제공이 가능함
- 개인신용평가모델 고도화를 통해 소득창출 능력과 변제 의지가 있는 개인을 대상으로 중금리 대출 지원에 기여할 예정임



화물차 개인사업자 신용도 평가모델 연구

한국도로공사

"안전하게 운전하는 화물차 개인사업자는 신용도에 가점을 줄 수 있을까요?"

화물차 개인사업자인 K씨는 최근 큰돈이 필요해 은행을 찾았다가 실의에 빠졌다. K씨의 수입이 정기적이지 않아 신용등급이 낮게 책정되었고, 이에 따라 필요한 만큼의 대출이 불가하다는 답변을 들었기 때문이다.

업종 상 수입이 비정기적인 거야 어쩔 수 없지만, 그래도 나처럼 교통법규를 준수하고 안전 운행하는 모범 운전자라면 지속적으로 일거리를 확보할 수 있다고 자부한다. 화물차 운행에서 수집되는 정보들을 신용도 평가에 반영하여 안전운행을 한다면 신용도에 가점을 줄 수는 없는 걸까?



1. 추진배경 및 필요성

- 부정기적인 수입으로 신용점수가 저평가되는 화물차 개인사업자의 안전운행 여부 및 고속도로 이용 횟수, 통행료 납부 정보 등을 개인 신용도 평가에 반영하여 모범 화물차 운전자의 금융 서비스 기회를 확대하고자 함
- 화물차 개인사업자는 안전운행 시 신용점수 향상을 통해 금융서비스 기회가 확대되고, 자발적인 안전운전을 유도하여 화물차 미납통행료 및 교통사고 감소의 긍정적 효과가 기대됨

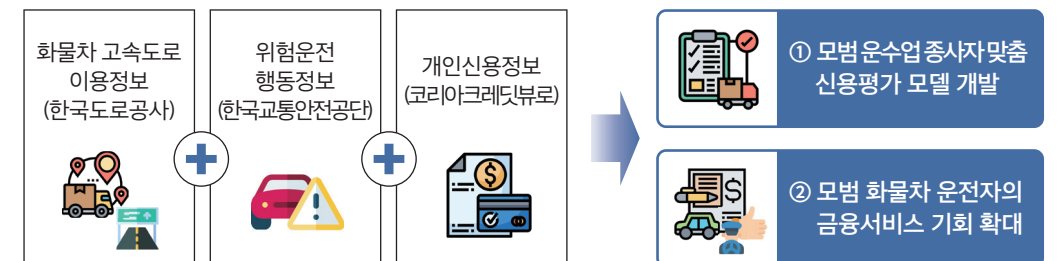
2. 결합데이터

	데이터 보유기관	데이터 개요	레코드 수	칼럼 수	데이터크기
1차 결합	한국도로공사	고속도로 이용 화물차 차량번호, 통행횟수, 완납 / 미납요금 등	295,337개	7개	9MB
	코리아 크레딧뷰로(주)	차량번호 보유 고객에 대한 신용평가 정보	7,980,573개	16개	488MB
	한국교통 안전공단	화물차량의 과속, 급가감속, 급회전 등 위험운전행동 횟수	18,906개	26개	5MB
2차 결합	한국도로공사	고속도로 이용 화물차 통행, 교통사고, 과적 / 적재불량 정보 등	382,102개	154개	166MB
	코리아 크레딧뷰로(주)	전국민의 대출, 연체, 카드영역의 신용정보 등	54,359,321개	204개	67GB

▶ 결합기관: 금융보안원

3. 분석결과 및 향후 계획

- 교통·안전정보를 활용한 신용평가 서비스를 개발하여 개인 신용평가모델을 고도화할 예정임
- 유관기관 간 제도 및 시스템을 연계·구축하여 교통·안전정보를 활용한 대국민 신용평가서비스를 제공할 예정임



가명정보 결합
아이디어

네트워크 회복을 통한 지역경제 활성화와 청년창업 지원

RE:DATA project
(경진대회 수상팀)

"카페 창업을 하려는데 입지 선정이 어려워요"

카페 창업을 준비중인 I씨는 요즘 길거리를 걸을 때마다 손님이 북적이는 가게들의 입지를 분석 해본다. 이미 유동인구가 많은 지역에 카페를 열었다가 한차례 실패를 맛봤기 때문이다. 단순히 유동인구가 많은 지역이라고 해서 그들이 모두 내 카페로 유입되는 것은 아니라는 것을 깨달을 수 있던 뼈아픈 경험이었다.

특정 장소를 지나가는 사람들의 연령, 성별, 방문목적, 방문 시간대를 알 수 있다면 그에 맞춰 적당한 위치를 선정하고, 그들의 취향과 특성에 맞는 메뉴로 손님들의 발길을 사로잡을 수 있을텐데 그걸 알아내기가 너무 어렵다.



1. 추진배경 및 필요성

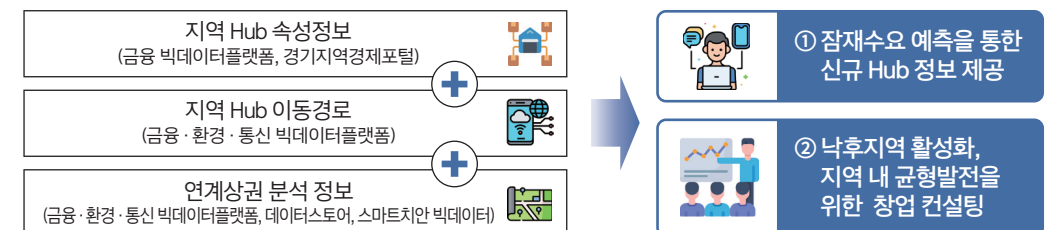
- 코로나19로 인해 침체된 지역상권 회복과 창업 성공률 향상을 위해 데이터 기반 컨설팅의 필요성이 대두됨
- 온라인으로 경험하기 어려운 문화 · 체험시설을 지역별로 제안하여 유동인구 유입방안을 모색하고 대면영업을 활성화하고자 함
- 공공사업 추진이 어려운 소규모 유동인구 집중시설(이하 '지역 Hub')에 대한 데이터를 제공하여 상권 활성화 및 업종 간 상생구조를 유도하고자 함

2. 결합데이터

분류	데이터명	제공기관
지역Hub 속성	· 가맹점 업종 / 지역 & 방문고객 주소, 성별, 연령 매출집계 · 가맹점별 성별, 연령별 매출 비율 데이터 · 가맹점별 요일, 시간 매출 비율 데이터 · 개별 상점별 고객 충성도 분석 · 국토도시 계획개발	금융 빅데이터 플랫폼 경기지역경제포털
지역Hub 이동 경로 분석	· 대중교통이용 정보 · 행정동별 보행 인구 · 시간 - 지역별, 업종 - 지역별 배달 주문건수	환경 빅데이터 플랫폼 금융 빅데이터 플랫폼 통신 빅데이터 플랫폼
연계상권 분석	· 행정동별 유동 / 생활 인구 · 전국 아파트 / 동별 세대 특성 통계정보 · 유동 인구 내역 · 교통 및 유동인구량 · 매출 기반 상권분석 데이터 · 임대료	금융 빅데이터 플랫폼 교통 데이터 거래소 통신 빅데이터 플랫폼 데이터스토어 스마트치안 빅데이터

3. 제안내용 및 기대 효과

- 골목 및 신규상권 중심으로 지역Hub 유형을 분류하고, 지역Hub의 특징과 방문고객 속성을 기반으로 연계상권에 대한 잠재고객을 예측함
- 신규Hub와 도시 개발계획을 기반으로 상권발전 유망지역을 예측하고, 잠재수요 예측 데이터와 가망 지역 예측 데이터를 통합하여 지역 · 업종별 맞춤형 추천 서비스를 제공함
- 누구나 쉽게 이해할 수 있는 데이터 분석 결과를 제공하여 데이터에 기반한 창업 및 의사결정 문화가 확산될 것으로 기대됨
- 정부 정책 중심의 일방향성을 탈피하여 민관합동의 전략을 수립하고, 지속적인 창업지원으로 낙후 지역 활성화 등 지역 내 균형발전에 기여할 수 있음



교통약자의 이동편의를 위한 대중교통 시설 점검 아이디어 도출

T-BTS
(경진대회 수상팀)

"교통약자가 자주 이용하는 역의 시설 먼저 점검해주세요"

2년 전 사고로 다리를 크게 다쳐 휠체어를 이용하는 L씨는 대중교통으로 이동해야 하는 상황이 생기면 걱정부터 앞선다. 계단 대신 리프트와 엘리베이터를 이용해야 하는데, 시설이 정상적으로 작동하지 않아 곤란했던 적이 종종 있었기 때문이다.

고장신고 접수 후 사후적으로 조치하는 방식에서 나아가, 교통약자가 많이 이용하는 정류소의 시설부터 선제적으로 점검할 수 있는 방법은 없을까? 그렇게 된다면 많은 교통약자의 안전을 보장하고 한정적인 예산도 효율적으로 집행할 수 있을 것 같다.



1. 추진배경 및 필요성

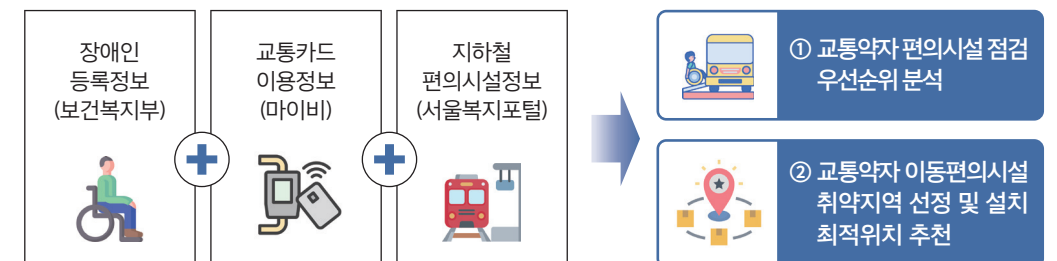
- 전동휠체어 이용자가 지난 10년 간 5배가량 증가했지만 열악한 교통약자 이동편의시설 탓에 교통약자는 크고 작은 사고의 위험에 노출되어 있음
- 현재는 이동편의시설이 노후되거나 훼손된 경우, 신고 접수 후 조치가 이루어지기 때문에 선제적인 시설점검의 필요성이 대두됨

2. 결합데이터

데이터명	제공기관
장애유형·정도별 등록 현황	보건복지부
교통카드 데이터	마이비
지하철 편의시설	서울복지포털

3. 제안내용 및 기대 효과

- 노후된 시설을 선제적으로 점검하여 교통약자의 편의성을 증대하고 사고를 예방할 수 있음
- 교통약자가 자주 이용하는 정류소부터 시설의 상태를 점검하고 편의시설을 구축하여 제한된 예산을 효과적으로 활용할 수 있음
- 교통약자가 자주 이용하는 데이터 정보를 활용하여 교통약자 대상 새로운 서비스를 개발하여 제공할 수 있음



보건의료데이터 통합과 직업력 결합을 통한 국민직업건강정보체계 구축

최강이 이김
(경진대회 수상팀)

"직업과 건강, 밀접한 관계가 있을 텐데 직업과 일생의 건강정보를 통합해서 관리할 순 없나요?"

대학생 P씨는 최근 흥미로운 기사를 읽었다. 머리카락만으로 어떤 일을 하는 사람인지 알아내는 실험에 관한 내용이었다. 생활하며 접하는 물질들이 몸속에 축적되면서 머리카락으로 기록이 남는다고 했다. 실제로 머리카락 분석만으로 도색업자, 용접공, 어부 등의 실험 참가자의 직업을 모두 맞춰냈다.

이렇게 머리카락에 근로환경의 흔적이 남는다면 분명 근로환경과 건강은 밀접한 관계가 있을텐데, 한 사람의 출생부터의 건강정보와 직업정보를 결합한다면 직업별 건강관리체계를 제시하고 데이터 기반의 보건의료서비스가 가능하지 않을까?



1. 추진배경 및 필요성

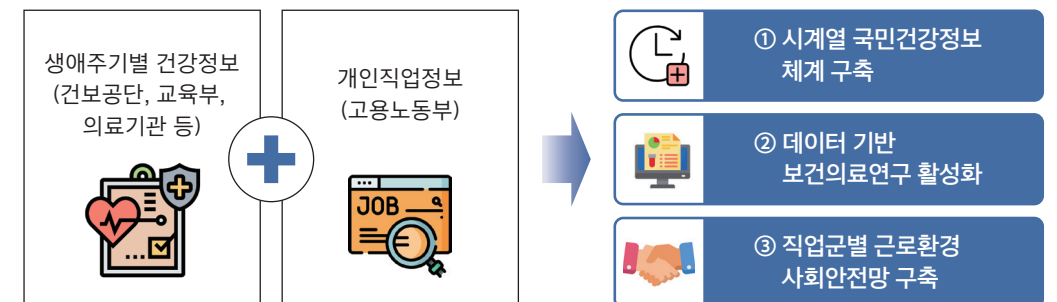
- 저출산·고령화로 인한 인구구조 변화 및 4차 산업혁명으로 인한 산업구조 변화로 고령층 및 근로자 등에서 새로운 형태의 보건의료 수요가 발생함
- 개인의 건강에 영향을 미치는 다양한 요인 중 근로 환경의 중요성이 대두되고 있음

2. 결합데이터

데이터명	제공기관
영유, 학교 밖 청소년, 일반건강검진 결과 데이터	국민건강보험공단
초·중·고등학교 학년별 건강검사(체중, 키 등) 데이터	교육부
병원에서 발생하는 모든 데이터(진료, 진단, 입원, 수술 등)	각 의료기관
예방접종 관련 데이터(시기, 종류 등)	질병관리청
진료비 산정을 위한 환자의 진료내역 데이터	건강보험심사평가원
개인단위의 일자리 정보 데이터	고용노동부

3. 제안내용 및 기대효과

- 개인의 전 생애에 걸쳐 축적된 건강정보를 통해 기존의 결과 중심 건강지표의 한계를 극복하고 종합적인 국민건강정보 통계를 생성할 수 있음
- 보건의료 분야 연구의 핵심과제인 양질의 데이터 확보 문제를 개선하여 보다 다양하고 과학적인 연구 설계가 가능해짐
- 특정 근로환경에서 발생 가능성이 큰 질병과 위험요인을 미리 파악하고 대안을 마련하는 등 사회적 안전망을 구축할 수 있음



선택과목별 학생 성취도 대조분석을 통한 효율적인 선택과목 운영 아이디어 도출

최형빈
(경진대회 수상팀)

"고교학점제* 도입, 효율적으로 선택과목을 운영할 수 있을까요?"

2025년부터 고교학점제가 모든 일반계 고등학교에 시행된다. 초등학교 고학년생 딸을 둔 S씨는 아이가 고등학교에 진학하면 대학생처럼 자신이 원하는 과목을 골라 들을 수 있다니 아이의 진로 개발에 큰 도움이 될 것 같아 기대에 부풀다.

하지만 한편으로는 각 학교에서 여러 선택과목 간 성적편차나 평가지표 등을 어떻게 표준화하여 운영할 수 있을지 궁금증이 든다. 학생 맞춤형 교육을 원활하게 운영하기 위해서는 이러한 기준이 먼저 필요할 것 같다.



* 고교학점제 : 학생이 기초 소양과 기본 학력을 바탕으로 진로·적성에 따라 과목을 선택하고, 이수기준에 도달한 과목에 대해 학점을 취득·누적하여 졸업하는 제도(2021, 교육부)

1. 추진배경 및 필요성

- 2025년까지 전국의 모든 고등학교에 고교학점제가 도입되면서 학생들은 자신이 원하는 다양한 과목을 직접 선택해 수강하고 학점 기준을 충족하여 졸업하게 됨
- 현재는 고교학점제로 인해 각 학교에서 다양하게 개설되는 선택과목들의 운영성과와 각 과목을 수강하는 학생들의 성취도 수준을 파악할 수 있는 수단이 부족함

2. 결합데이터

데이터명	제공기관
학생 선택과목 선택 데이터	교육부 및 전국 17개 시도교육청
학생 선택과목 성적, 성취도, 출석 등 데이터	

3. 제안내용 및 기대효과

- 학생들의 선택과목 선택 데이터와 수강중인 선택과목 수업의 성적, 성취도, 출석률 등을 결합하면 각 선택과목별, 각 학교별 운영 편차를 분석할 수 있게 됨
- 선택과목의 교육 수준을 동일 학교 내 타 선택과목 및 타 학교의 동일 선택과목 등과 비교하여 해당 학교에서 운영하는 선택과목의 교육수준을 파악할 수 있음
- 각 선택과목의 담당교사의 교육수준 및 교육 편성 등에 따라 각 학교별 선택과목의 교육 수준의 편차 발생이 불가피하므로, 가명결합 데이터 분석을 통해 각 학교는 선택과목별 교육수준을 파악하여 대안을 마련함으로써 운영 효율성 및 교육의 질을 제고할 수 있음



2021 가명정보 활용 우수사례집

발 행 일 2022년 2월

발 행 처 개인정보보호위원회, 과학기술정보통신부, 금융위원회, 보건복지부

지원기관 한국지능정보사회진흥원

본 사례집 내용의 무단 전재를 금하며,
가공·인용 시에는 출처를 밝혀 주시기 바랍니다.