

금융취약지 특성 분석 및 금융소외 해소방안 마련

분석 결과보고서

소 속	직 위	이 름
행정안전부	일경험 수련생	김지영(Y220699)
		전지우
우체국금융개발원	전임연구원	박명지

목 차(예1)

I. 분석 개요	1
1. 배경 및 필요성	1
2. 수행 기간	1
3. 분석 목표	3
4. 수행 체계	3
II. 분석 방법	4
1. 분석 프로세스	4
2. 분석 도구/환경	5
3. 활용 데이터	5
4. 분석 방법	7
III. 분석 결과	9
1. 수행 결과	9
IV. 결론	27

목 차

1. 분석 개요	4
1.1. 분석 배경 및 개요	
1.1.1 분석 배경	
1.2. 분석 목적 및 방향	
1.2.1 현황 탐색	
1.2.2 분석 방향	
2. 분석 데이터	6
2.1. 분석 데이터 목록	
2.2. 데이터 상세 설명	
2.3. 데이터 정제 방안	
3. 분석 프로세스	12
3.1. 분석 프로세스	
3.2. 분석 내용 및 방법	
3.2.1 데이터 수집 및 전처리	
3.2.2 EDA	
3.2.3 금융취약지수 정의 및 산출	
3.2.4 EDA	
3.2.5 데이터 분석 모델링	
4. 분석결과	29
4.1 금융취약지 특성 변수 T-검정	
4.2 Random Forest 기반 특성 중요도 산출	
4.3 중요도 상위변수 차원축소	

4.4 금융취약지 군집화 및 결과도출	
5. 금융소외 해소방안 및 기대효과	44
6. 참고문헌(Reference)	46

1. 분석 개요

1.1 분석 배경 및 개요

1.1.1 분석 배경

금융 산업의 디지털 전환이 가속화됨에 따라 은행의 고객 접점은 모바일 비대면 채널로 집중되고 있다. 실제로 연간 대면채널은 연평균 3.43%로 감소한 반면에 모바일 뱅킹앱을 설치하고 이용하는 활성화기기는 12.86%로 증가했다. 금융당국은 점포폐쇄 관련하여 공동절차를 마련하는 등 급격한 점포 폐쇄로 인한 금융소외를 방지하고자 다방면으로 정책을 마련하고 있으나 운영 효율성과 수익성의 관점에서 은행의 점포폐쇄 결정 자체에 대한 직접적인 개입은 불가하다. 따라서 적극적인 사전 대응을 통한 금융소외 해소 효과는 미비한 상황이며 공익적 차원의 정책·제도 마련이 필수 불가결한 상황이다. 따라서 금융접근성에 대한 정의와 함께 금융 취약지를 선별하여 찾고 공익적 관점에서 금융소외현상의 해소방안을 모색하고자 한다.

구분	2015년	2016년	2017년	2018년	2019년	2020년	2021년
전체은행(개)	7,281	7,101	6,789	6,766	6,709	6,405	6,094
4대은행(개)	3,924	3,757	3,575	3,563	3,525	3,303	3,276
뱅킹앱 활성화기(만 개)	-	-	-	6,169	7,240	8,262	8,868

1.2 분석 목적 및 방향

1.2.1 현황 탐색

금융산업의 디지털 전환이 전방위적으로 가속화됨에 따라 해가 거듭될수록 금융 소비자의 연령별·소득별·지역별 비대면 채널 및 서비스 이용격차가 커지고 있다. 대면채널 의존도가 높은 금융취약계층을 중심으로 금융거래 편의성이 크게 떨어지고 있다.

연령별 모바일 금융서비스 이용경험을 살펴보면 60대 이상의 연령층에서 5대 시중은행과 인터넷전문은행의 모바일 채널 이용도가 타 연령대비 크게 떨어지는 것을 볼 수 있다. 특히 인터넷전문은행의 경우 60대 이상의 연령층에서는 4.5%만 사용 중이며 모바일 채널을 통해 제공되는 우대금리, 모바일 특화상품, 제휴 마케팅 혜택 등에서 대부분 소외되었다고 볼 수 있다.

소득별 모바일 금융서비스 이용경험을 살펴보면 연소득이 높을수록 모바일을 통

한 금융서비스 이용도가 활발한 것을 확인 할 수 있다. 1인 가구 기준 연간 중위 소득이 약 2,300만원인 것을 고려했을 때 우리나라 평균적인 소득수준에 미치지 못하는 가구일수록 모바일 금융서비스 이용에서 소외된 것으로 보인다. 이는 모바일 채널 위주의 금융서비스 제공이 꼭 금융 접근성의 확대를 의미하는 것은 아니며 오히려 개인적 금융소외를 지속적으로 확대할 수 있다는 것을 시사한다.

도시지역이 아닌 농·어촌 거주자의 경우 이용 중인 디지털 서비스 유형을 살펴보면 금융거래는 응답자의 56.5%가 이용하고 있다. 농어촌의 특성상 고령인구가 상대적으로 더 많이 거주하고 있으며 대면 창구를 통한 금융거래가 보다 활발하기 때문에 이용률 뿐 만 아니라 일반 국민과의 이용률 격차 또한 큰 편이다.

연령별, 소득별, 거주지별로 살펴본 결과 연령이 높을수록, 소득이 낮을수록, 농·어촌 거주민일수록 모바일 채널을 통한 금융서비스 이용경험이 저조한 상황이며 따라서 금융취약계층을 대상으로 금융서비스를 제공할 경우 채널과 이용률 제고에 대한 일반국민과 차별화되는 방안 모색이 필요하다.



1.2.2 분석 방향

금융소외에 대해서는 크게 개인적 금융소외와 지리적 금융소외로 구분한다. 먼저 개인적 금융소외란 금융소비자의 소득·신용에 따라 제도권 금융상품 이용에의 접근성 결여를 의미하는 것으로 금융 상품 및 서비스 이용의 어려움을 의미한다. 다음으로 지리적 금융소외란 은행점포가 지역 내 부족 또는 부재하여 발생하는 접근성 결여로 거리적·물리적인 이동상의 어려움 때문에 발생하는 금융소외를 의미한다. 본 연구에서는 금융 접근성의 의미를 '지리적 접근성'으로 한정하여 살펴보고 금융 취약성을 지리적 접근성 결여에 따른 서비스 이용의 제한 또는 취약을 의미하는 것으로 한정한다. 또한 금융취약계층은 금융취약에 노출된 계층으로 대표적으로는 고령자, 장애인, 농·어촌 거주자로 한정하여 논의하고자 한다.

2. 분석 데이터

2.1. 분석 데이터 목록

데이터명	시점	유형	출처
장애인 인구수(시군구)	2021	csv	통계청
고령층 인구수(시군구)	2021	csv	통계청
전국 은행지점 데이터	2021, 2018, 2015년 12월	xlsx	은행연합회
전국 도로망 데이터	2020	shp	국가도로교통DB
사업체/종사자수(시군구)	2019	xlsx	통계청
우체국 관서 데이터	2021	csv	공공데이터포털
전국 행정구역 경계(시군구)	2021	shp	국가공간정보포털
행정동별 통계	2021	csv	신한카드_SKT_KCB
분기X행정동X라이프스테이지	2021	csv	신한카드_SKT_KCB

2.2. 데이터 상세 설명

□ 장애인 인구수(시군구)

통계청 전국 시군구별 장애인 인구 데이터로 지역별 장애인 인구수 정보를 제공.

1) 시군구별,장애등급별,성별 등록장애인수

『장애인현황』, 보건복지부 (자료문의처: 044-202-3290) 통계실명자료 윤리인간영문

수록기간: 년 2007 ~ 2018 / 자료갱신일: 2021-08-19 / 주석정보

시점, 기간/종류, 행정안전, 열고정보, 새 탭 열기, 화면복사, 주소/출처, 스크랩, 인쇄, 다운로드, 조희상

(단위: 명)

시군구별(1)	시군구별(2)	2018		1급			
		합계	남자	여자	합계	남자	여자
전국	소계	2,585,876	1,496,298	1,089,578	198,281	116,462	81,819
서울특별시	소계	392,920	227,910	165,010	32,730	19,553	13,177
	종로구	6,017	3,493	2,524	493	283	210
	중구	5,638	3,265	2,373	452	263	189

□ 고령층 인구수(시군구)

통계청 전국 시군구별 고령층 인구 데이터로 지역별 고령층 인구수 정보를 제공.

고령인구비율(시도/시/군/구)

통계청@통계청백과 (자료문의처: 042-481-3608)

수록기간: 월, 년 2000 ~ 2022.10 / 자료갱신일: 2022-11-03 / 주석정보

시점, 기간/종류, 행정안전, 열고정보, 새 탭 열기, 화면복사, 주소/출처, 스크랩, 인쇄, 다운로드, 조희상

2) 행정구역별(1)	2022.06		65세이상인구 (A) (명)	전체인구 (B) (명)	2022.07		65세이상인구 (A) (명)	전체인구 (B) (명)	2022.08		65세이상인구 (A) (명)
	고령인구비율 (A÷B×100) (%)	(%)			고령인구비율 (A÷B×100) (%)	(%)			고령인구비율 (A÷B×100) (%)	(%)	
충청북도	19.4	309,395	1,597,118	19.4	310,417	1,597,163	19.6	312			
충청남도	20.2	428,627	2,120,201	20.3	429,894	2,121,011	20.4	431			
제주특별자치도	16.7	112,991	678,012	16.7	113,407	678,491	16.8	114			

□ 전국 은행지점 데이터

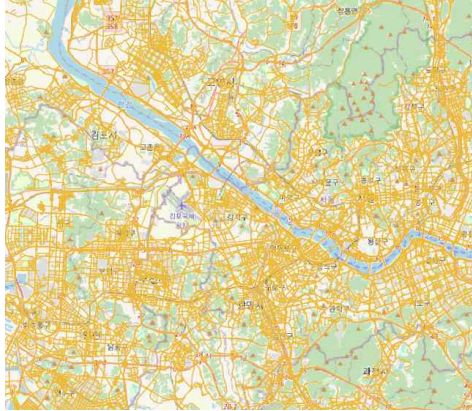
전국은행연합회가 반기별로 공시하는 자료로 해당연도(2015, 2018, 2021년) 12월말 기준 전국 시중은행, 상호금융, 지방은행 지점 정보를 제공.

□ '21.12월말 기준 유인점포 영업현황(반기보)

은행명(1)	지포코드(2)	점포명(3)	지점/종원소(4)	점포형태(대)(5)	점포형태(소)(6)	모형명(7)	영업본부명(8)	주소(9-14)					
								지/시/도(9)	구/시/군/구(10)	구/시/도(11)	반년별영점구역(12)	연말년도 동(13)	도포명(14)
신협은행	0027083	강남	지점	기점			강남지역본부	서울특별시	강남구	강동동		18-1	양동대로 508
신협은행	0027084	도곡	지점	기점			강남지역본부	서울특별시	강남구	도곡동		467-7	남부순환로 3303
신협은행	0027083	반포	지점	기점			강남지역본부	서울특별시	서초구	반포동		18-3	반포대로 287

□ 전국 도로망 데이터

국가교통DB에서 제공하는 전국 도로망 데이터로 전국 도로의 ID, 유형, 위치정보 등 상세현황을 공간 정보인 shp 파일 형식으로 제공. (Line 데이터)



□ 사업체/종사자수(시군구)

통계청 시도별 사업체와 종사자수 데이터로 지역별 사업체수와 종사자수를 제공한다.

1) 산업별 사업체수 및 종사자수 「경기도구리사기본통계」, 경기도 구리시 (자료문의처: 031-550-2082) 통계실명자료

수속기간: 년 2013 ~ 2020 / 자료갱신일: 2022-08-17 / 주석정보

시역, 읍면/읍면, 행정전환, 업고정해제, 새팀 열기, 화면복사, 주소/층지, 스크랩, 인쇄, 다운로드, 초화상

산업별(1)	산업별(2)	산업별(3)	2018	2019	2020
합계	사업체수 (개)	소계	14,648	14,826	21,866
		여성 대표자	5,733	5,742	8,059
	종사자수 (명)	계	66,759	66,383	68,787
		남	33,048	33,128	36,107
		여	33,711	33,255	32,680

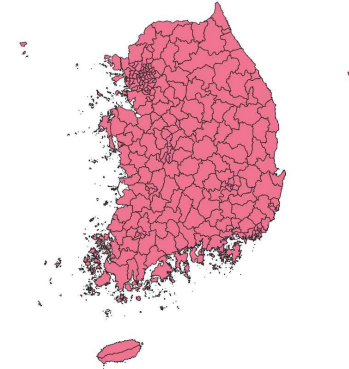
□ 우체국 관서 데이터

전국에 위치한 우체국 관서 중 금융업무를 취급하는 관서의 정보를 포함한다.

지방우정청	출발국명	우체국명	주소(지번)	주소(도로명)	전화번호
서울지방우정청	서울중앙우체국	서울중앙우체국	서울특별시 중구 중무로1가 21-1	서울특별시 중구 소공로 70 (중무로1가)	026450-1114
서울지방우정청	서울중앙우체국	서울중앙우체국	서울특별시 중구 을지로4가 312-1	서울특별시 중구 을지로 154-2 (을지로4가)	022273-6505
서울지방우정청	서울중앙우체국	서울중앙우체국	서울특별시 중구 을지로 42 우익빌딩 2층	서울특별시 중구 을지로 42 우익빌딩 2층 (을지로4가)	023783-7025
서울지방우정청	서울중앙우체국	서울중앙우체국	서울특별시 중구 남산로 190-11	서울특별시 중구 남산로 190 (남산동)	02756-3377
서울지방우정청	서울중앙우체국	서울중앙우체국	서울특별시 중구 다산로29길 23(신당동)	서울특별시 중구 다산로 29길 23 (신당동)	022252-5365

□ 전국 행정구역 경계

시군구 단위 전국 행정구역 경계를 shp파일 형태로 제공. QGIS를 통해 금융취약지수 계산과 시각화에 이용한다.



□ 행정동별 통계

외부 데이터로 반출 불가능해 클라우드 환경 내에서 분석에 이용하였다. 신한카드_SKT_KCB의 데이터로 총 56개 열(변수)로 구성되어 행정동별 소득, 자산 및 관심사 및 선호도 등의 통계를 보여준다.

행정동코드	고객수	하이엔드_소비수준	하이엔드_소득수준	하이엔드_법인대표/전문직	주정연소득	연간카드소비금액	순자산평가금액	최근차량발행부담액	국산차량보유	수입차량보유	20대 대상자
11110515	1898	138	143	29	85256390	52312000	719864440	8632119	135	31	270
11110530	1536	196	275	56	90270820	56203900	947738550	7682153	88	34	192
11110540	379	32	23	10	16357290	11623900	151632160	1860718	29	7	55
11110550	1156	107	124	19	53967930	37725000	434289670	6319766	106	24	145
11110560	2438	313	339	66	133607520	91011600	1271698470	16434023	214	96	269

□ 분기X행정동X라이프스타일

외부 데이터로 반출 불가능해 클라우드 환경 내에서 분석에 이용하였다. 신한카드_SKT_KCB의 데이터로 총 61개 열(변수)로 구성되어 2020년 분기 및 행정동별 업종 및 고객유형에 따른 소비규모 데이터를 보여준다.

기간시점	행정동코드	라이프스타일	요식이용금액	유흥이용금액	백화점이용금액	대형마트이용금액	소형유통점이용금액	의류/장화이용금액	스포츠/문화/레저이용금액	숙박이용금액	여방이용금액
201Q	11110515	4	91583000	2219000	23012000	21918000	121767000	13647000	20132000	2895000	8749000
201Q	11110515	2	63029000	1484000	9638000	10451000	58278000	8788000	8595000	1804000	1568000
201Q	11110515	1	176909000	5659000	11535000	17639000	107835000	16570000	26224000	6891000	6039000
201Q	11110515	5	58203000	632000	13509000	11936000	59150000	7138000	16287000	4891000	2648000
201Q	11110515	3	158106000	6873000	46979000	33238000	171461000	17625000	28917000	2976000	7960000

2.3. 데이터 정제 방안

□ 장애인 인구수(시군구)

Excel에서 시군구별 전처리를 통해 시군구별 장애인 수의 합계 형태로 된 csv 파일을 생성한다.

□ 고령층 인구수(시군구)

Excel에서 시군구별 전처리를 통해 시군구별 고령층 수의 합계 형태로 된 csv 파일을 생성한다.

□ 전국 은행지점 데이터

전국은행연합회 데이터를 바탕으로 도로명 주소 파생 칼럼을 생성한다. 이후 Geocoding을 통해 도로명 주소를 기준으로 위·경도를 얻는다.

□ 전국 도로망 데이터

전국 도로망 shp 형식 데이터로 QGIS를 이용한 OD Matrix 분석 시 거리 측정
에 사용한다.

□ 사업체/종사자수(시군구)

Excel에서 시군구별 전처리를 통해 시군구 당 사업체 수와 종사자수 합계의 형태로 된 csv 파일을 생성한다.

□ 우체국 관서 데이터

전국에 위치한 우체국 관서 중 금융취급 관서만 전처리한 뒤 도로명 주소를 기반으로 Geocoding 과정을 거쳐 위경도를 얻는다.

□ 전국 행정구역 경계

전국 행정구역 경계 shp 형식 데이터로 QGIS를 이용한 OD Matrix 분석 시 거리 측정에 사용한다.

□ 행정동별 통계

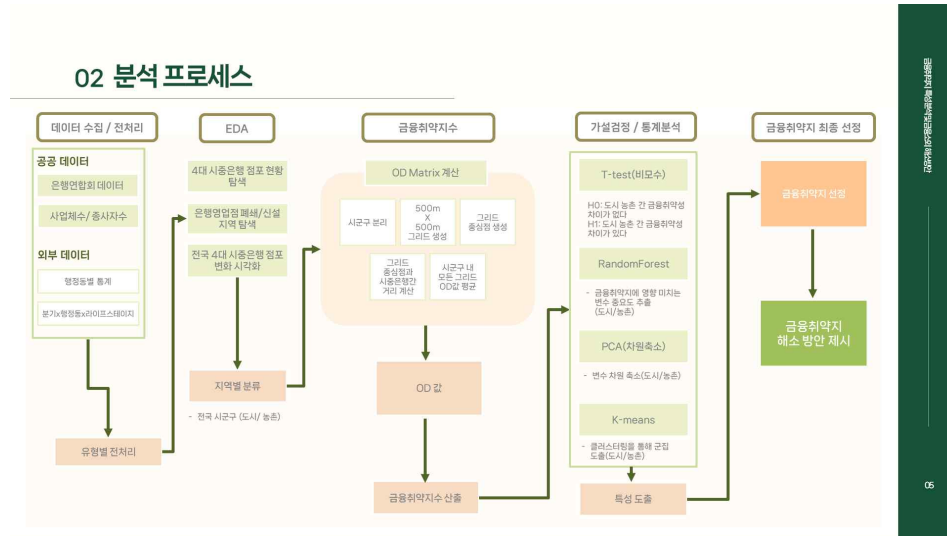
행정동 단위의 번듯값을 합하여 시군구 단위 데이터로 변환한다. 이때, “강서구(서울특별시/부산광역시)”, “고성군(강원도, 경상남도)”와 같이 지명이 중복되는 경우를 고려하여 “시도+시군구”로 행정구역명을 지정한다.

□ 분기X행정동X라이프스타일

2020년 4분기에 해당하는 데이터를 추출하고 행정동 단위의 번듯값을 합하여 시군구 단위 데이터로 변환한다. 이때, 지명이 중복되는 경우를 고려하여 “시도+시군구”로 행정구역명을 지정한다.

3. 분석 프로세스

3.1. 분석 프로세스



- 분석 과정은 데이터 수집 및 전처리 → EDA → 금융취약지수 도출 → 가설검정 및 통계분석 → 금융취약지 최종 선정 순으로 이루어진다.

3.2. 분석 내용 및 방법

3.2.1 데이터 수집 및 전처리

분석을 위해 사용한 데이터는 다음과 같다. 공공데이터로는 은행연합회 데이터와 통계청의 사업체 수/종사자 수 데이터를 사용하였다. 외부 데이터로는 공모전 주최 측에서 제공하는 신한카드_SKT_KCB 데이터를 사용하였다. 외부데이터와 내부데이터를 각각 전처리·가공하여 분석에 활용하였다.

□ 은행연합회 데이터

은행연합회 데이터의 경우, 4대 시중은행을 기준으로 데이터를 추출하였다. 본 보고서에서의 금융취약지 분석은 상호금융을 제외한 4대 시중은행을 기준으로 하기 때문에 전체 데이터에서 4대 시중은행만 추출하여 분석에 활용하였다. 추출된 은행연합회 데이터는 카카오 API를 이용하여 Geocoding 과정을 거쳐 위·경도를 추출한 이후 EDA와 OD Matrix 분석 과정에 이용하였다.

■ 데이터 정제 과정

- ① 은행연합회 데이터(2015, 2018, 2021) 중 4대 시중은행 추출
- ② 추출된 은행연합회 데이터에 주소 파생 칼럼 생성
- ③ 생성된 주소를 기준으로 카카오 API를 이용한 지오코딩을 통해 위·경도 추출
- ④ 결측값의 경우 값 대체
 - * 위도가 존재하지 않는 경우 → 인근 위·경도로 대체
 - * 주소가 존재하지 않는 경우 → 도로명주소가 아닌 지번 주소를 사용하여 인근 위·경도로 대체

□ 신한카드_SKT_KCB 데이터

신한카드, KCB, SKT 3사의 결합가명 정보기반으로 구성된 고객 및 시장분석용 데이터셋이다. 신한카드, SKT, KCB 3사 가명결합된 670만명을 기준으로 거주지 기준 행정동별 데이터를 집계한 통계 데이터이다. 외부데이터에 해당하며 총 2개의 csv 파일로 제공되었다. 이 두 데이터셋에 들어있는 변수를 활용해 파생변수를 만들어 분석하였다.

■ dong_table1.csv : 행정동별 통계 데이터(20년말 시점 데이터)

원본 데이터(56개 변수) 중 8개 변수를 추출했다. 해당되는 변수로는 행정동코드, 하이엔드_소비대상자수, 하이엔드_소득대상자수, 추정연소득, 연간카드소비금액, 순자산평가금액, 국산차량보유자수, 수입차량보유자수 가 있다. 이중 추정연소득, 연간카드소비금액, 순자산평가금액은 계산의 편의를 위해 천만원 단위로 재설정하였다.

■ dong_table2.csv : 행정동별X라이프스테이지별 통계데이터(20년 분기별 4개 시점)

원본 데이터(61개 변수) 중 22개 변수를 추출했다. 해당되는 변수로는 지하철이용횟수, 버스 주중이용건수, 택시주중이용건수, 버스주말이용건수, 택시주말이용건수, 증권사정보, 암호화폐정보, 자동차수입정보, 자동차전기차 정보, 하이엔드명품 정보, 하이엔드백화점 정보가 있다.

원본 데이터에서 추출된 변수를 바탕으로 5개의 파생 변수가 생성되었다. 5개의 파생 변수로는 대중교통_주중, 대중교통_주말, 투자_정보, 자동차_정보, 하이엔드 소비_정보가 있다.

지하철 이용횟수 변수의 경우 주말과 주중의 구분이 없는 데이터이다. 따라서 이를 해결하기 위해 주중과 주말의 이용횟수를 0.75와 0.25로 나누어 가중치를 주었다(실제 지하철 요일별 이용횟수의 비율 참고). 이렇게 가중치가 적용된 지하철 이용횟수는 대중교통_주중, 대중교통_주말 파생 변수로 사용된다.

□ 파생변수 목록

- 대중교통_주중(주중 대중교통 이용량) : 지하철주중이용횟수(지하철이용횟수 * 0.75) + 버스주중이용건수 + 택시주중이용건수
- 대중교통_주말(주말 대중교통 이용량) : 지하철주말이용횟수(지하철이용횟수 * 0.25) + 버스주말이용건수 + 택시주말이용건수
- 투자_정보(금융투자에 대한 관심도) : 증권사_정보 + 암호화폐_정보
- 자동차_정보(자동차에 대한 관심도) : 자동차수입_정보 + 자동차전기차_정보
- 하이엔드소비_정보(고급 브랜드 소비에 대한 관심도) : 하이엔드명품_정보 + 하이엔드백화점_정보

□ 시군구 컬럼 조인

dong_table1.csv 와 dong_table2.csv의 데이터를 행정동 코드를 기준으로 조인한다. 데이터를 조인한 뒤 시도별 컬럼과 시군구명 컬럼을 더해 시군구_구분이라는 이름으로 파생 컬럼을 생성한다.

□ 최종 변수

분석에 사용할 최종 변수는 dong_table1.csv의 변수 + dong_table2.csv의 변수 + 시군구_구분 컬럼의 형태이다. 파생 변수들을 전처리 후 통합한 테이블의 각 컬럼을 분석변수로 사용한다.

□ OD-Matrix 이상치 제외

은행연합회 데이터를 바탕으로 QGIS에서 QNEAT3 플러그인을 이용해 OD Matrix를 계산한다. 그러나 울릉군의 경우 도서지역으로 육지에 위치한 은행 지점과의 거리가 멀어 OD Matrix 값이 NULL로 나오는 등 값이 제대로 측정되지 않는 결과가 발생했다. 따라서 울릉군의 경우 본 분석에서는 제외한다.

3.2.2 EDA - 시중은행 지역별 점포 현황

□ 4대 시중은행 점포 현황(12월 기준)



금융취약지 분석에 앞서 시중은행의 점포 폐쇄 현상이 금융취약지를 판단하는 기준이 될 수 있다고 판단했다. 따라서 시중은행의 점포 폐쇄 지역 특성 분석을 위해 은행연합회의 데이터를 기준으로 점포 폐쇄 현황을 살펴보았다. 2011년부터 2021년까지의 은행연합회 데이터를 통해 시각화하였다. 4대 시중은행 점포는 2012년 이후로 줄어드는 추세이며, 2015년에 3928개 지점으로 증가했다가 다시 줄어드는 추세를 보인다. 2014~2015년 사이 시중은행 점포 수가 증가하는 데에는 하나은행과 외환은행의 통합(2015년 9월 1일)이 큰 원인으로 작용한 것으로 보인다.

※ (참고) KEB 하나은행 영업점 감소량

2015 → 2016 : 72개 지점

2016 → 2017 : 87개 지점

2017 → 2018 : 22개 지점

□ 2015 ~ 2018년 4대 시중은행 영업점 변화

전국 4대 시중은행 영업점 변화를 지역에 따라 차이가 있다고 가정하고 분석하기 위해 전국을 시군구로 나눈다. 전국 시군구별 은행 점포 개수의 변화를 통해 점포 폐쇄 및 신설 지역을 탐색한다. 기준년도를 2015년, 2018년, 2021년으로 잡고 분석을 진행한다. 하나은행의 경우 점포통합으로 인해 타 시중은행과는 다른 양상을 보인다고 판단하여 점포통합이 일어난 해인 2015년 ~2018년 사이 데이터는 제외하고 분석했다.

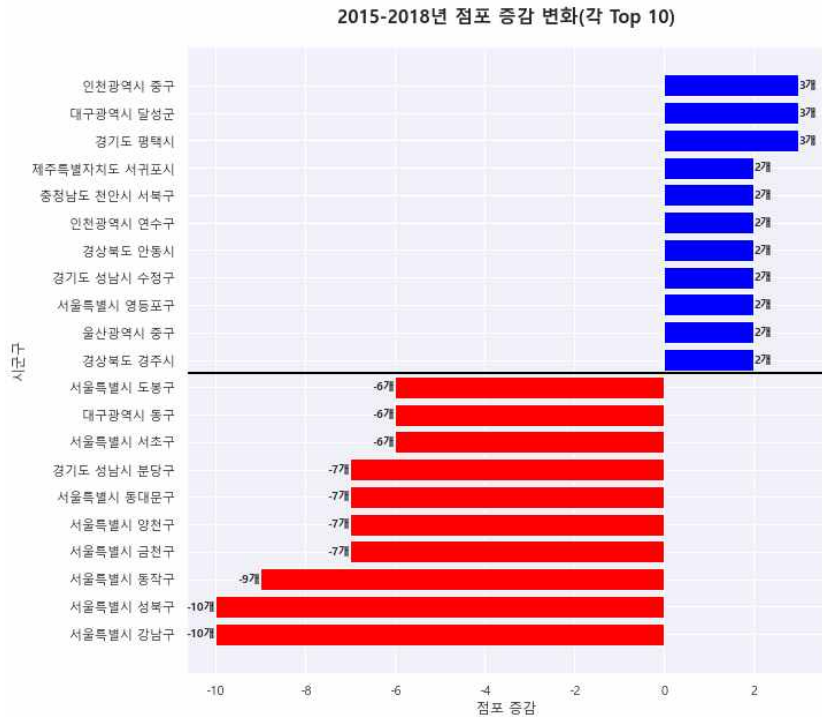
- 비교 대상: 국민은행, 신한은행, 우리은행, NH농협은행
- 기준: 광역시+구/군, 일반시+구/군 (*구가 없는 경우 시까지 적용)
- 데이터: 은행연합회 반기별(12월) 2015, 2018, 2021년 데이터

□ 행정구역 변화

행정구역 개편으로 인해 2015년 ~ 2021년 사이의 행정구역이 통합/변경되었다. 본 분석에서는 2021년 새로 변경된 행정구역을 기준으로 한다.

- 경기도 부천시 : 2016년 원미구, 소사구, 오정구 폐지 & 광역동으로 개편
→ 2015년, 2021년 데이터 “부천시” 단일 통일
- 인천광역시 미추홀구 : 본래 남구였으나 2018년 “미추홀구”로 명칭변경
→ 2015년, 2018년 데이터 “남구” ⇒ “미추홀구” 변경
- 충청북도 청원군 : 2014년 충청북도 청주시와 통합되면서 청원구, 상당구, 흥덕구, 서원구로 편입
→ 2015년, 2018년, 2021년 데이터 “청원군” ⇒ “청주시” 변경

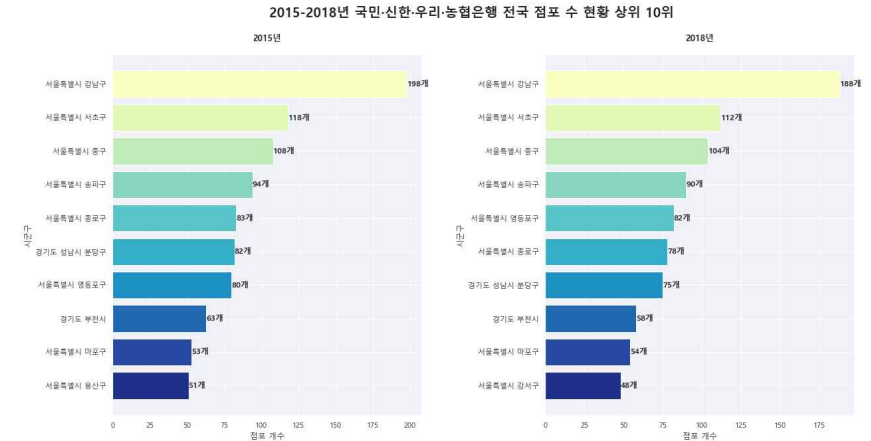
□ 2015 ~ 2018년 4대 은행 점포 증감 지역



2015년~2018년 은행 점포 감소량 상위 10개 지역	2015년~2018년 은행 점포 증가량 상위 10개 지역
서울특별시 강남구, 서울특별시 성북구, 서울특별시 동작구, 서울특별시 금천구, 서울특별시 양천구, 서울특별시 동대문구, 경기도 성남시 분당구, 서울특별시 서초구, 서울특별시 도봉구	인천광역시 중구, 대구광역시 달성군, 경기도 평택시, 제주특별자치도 서귀포시, 충청남도 천안시 서북구, 인천광역시 연수구, 경상북도 안동시, 경기도 성남시 수정구, 서울특별시 영등포구, 울산광역시 중구, 경상북도 경주시

시중은행은 은행 점포가 높은 밀도로 존재하고 있는 수도권과 광역시를 중심으로 급격하게 점포를 폐쇄하는 경향이 있으며, 인구가 증가하고 개발이 이루어지는 신도시 인근 지역 위주로 은행 점포가 늘어나는 경향을 띤다. 최대 10지점까지 점포를 축소하는 반면 점포 확대는 최대 3지점으로 점포 축소에 비해 확대 폭이 작다.

□ 2015 ~ 2018년 4대 은행 점포 수 상위 10위



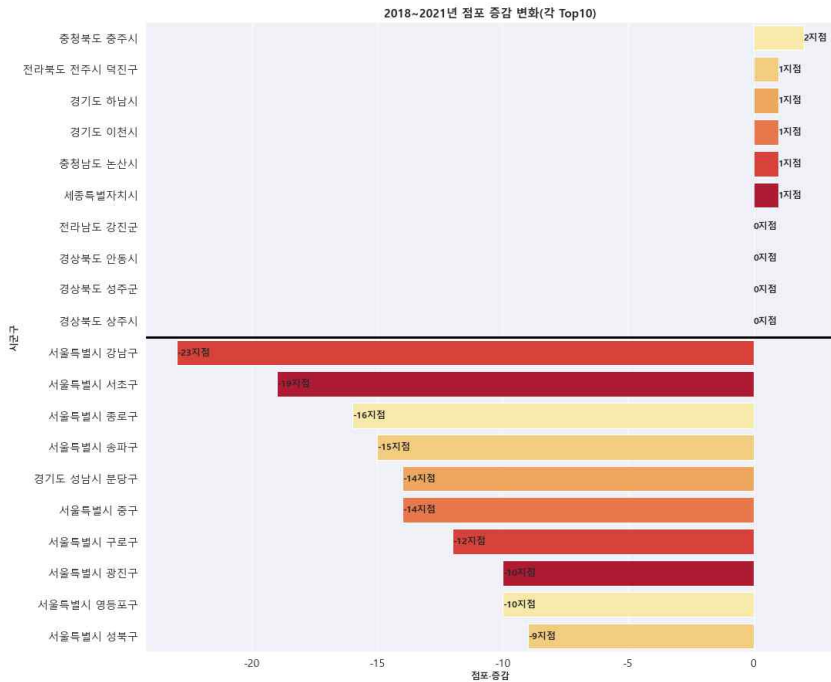
2015년 은행 점포 상위 10개 지역	2018년 은행 점포 상위 10개 지역
서울특별시 강남구, 서울특별시 서초구, 서울특별시 중구, 서울특별시 송파구, 서울특별시 종로구, 경기도 성남시 분당구, 서울특별시 영등포구, 경기도 부천시, 서울특별시 마포구, 서울특별시 용산구	서울특별시 강남구, 서울특별시 서초구, 서울특별시 중구, 서울특별시 송파구, 서울특별시 영등포구, 서울특별시 종로구, 경기도 성남시 분당구, 경기도 부천시, 서울특별시 마포구, 서울특별시 강서구

2015년과 2018년 모두 수도권, 도시 지역이 전국에서 점포 수 현황 상위순위를 차지했다. 시중은행의 경우에는 인구가 많은 도시 지역을 위주로 은행 점포가 많이 입지하는 것을 확인할 수 있다.

2018년에도 2015년과 마찬가지로 같은 비율로 은행 점포가 입지한다. 그러나 2015년에 198개의 지점이 위치했던 강남구는 2018년도에는 188개 지점으로 점포가 감소했고, 2위 지역인 서초구도 118개 지점에서 112개 지점으로 점포가 감소했다.

이를 통해 광역시, 수도권 등 대도시 지역 위주로 은행 점포가 입지하는 경향을 띤다고 볼 수 있으며, 은행은 점포 운영의 효율성을 위해 점포가 많은 지역의 경우 비용 절감을 위해 점포를 빠르게 축소해나가고 있다.

□ 2018 ~ 2021년 4대 은행 점포 증감 지역



2018년~2021년 은행 점포 감소량 상위 10개 지역	2018년~2021년 은행 점포 증가량 상위 10개 지역
서울특별시 강남구, 서울특별시 서초구, 서울특별시 종로구, 서울특별시 송파구, 경기도 성남시 분당구, 서울특별시 중구, 서울특별시 구로구, 서울특별시 광진구, 서울특별시 영등포구, 서울특별시 성북구	충청북도 충주시, 전라북도 전주시 덕진구, 경기도 하남시, 경기도 이천시, 충청남도 논산시, 세종특별자치시

2018년과 마찬가지로 은행 점포가 높은 밀도로 존재하고 있는 수도권과 광역시를 중심으로 점포를 폐쇄하는 경향이 뚜렷하다. 인구가 증가하고 개발이 일어나는 신도시 지역 중심으로 점포가 증가하나 증가 폭이 크지 않다. 점포 축소는 최대 23개 지점까지 축소하는 데 비해 점포 확대는 최대 2개 지점으로 점포 축소에 비해 확대 폭이 작다.

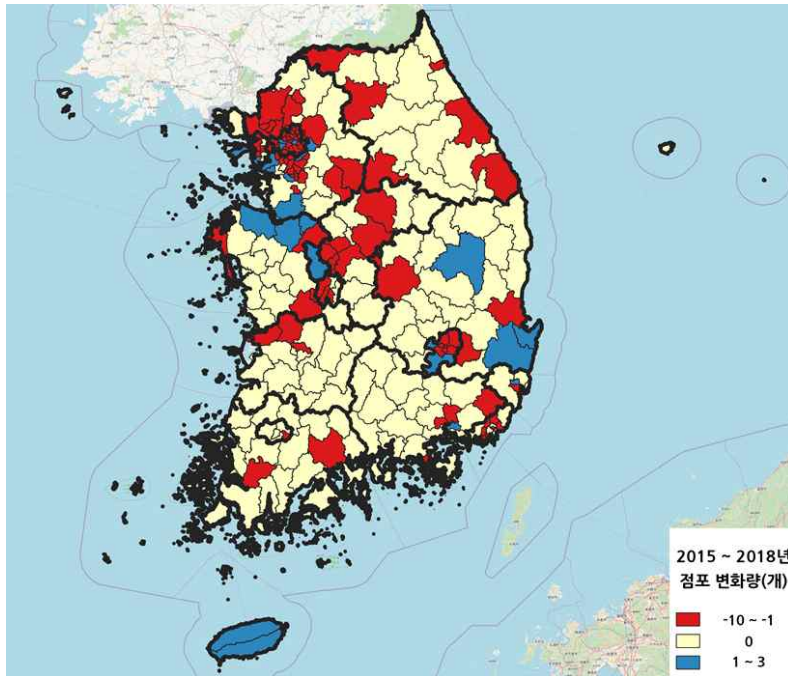
□ 2018 ~ 2021년 4대 은행 점포 수 상위 10위



2018년~2021년 은행 점포 감소량 상위 10개 지역	2018년~2021년 은행 점포 증가량 상위 10개 지역
서울특별시 강남구, 서울특별시 서초구, 서울특별시 종로구, 서울특별시 송파구, 경기도 성남시 분당구, 서울특별시 중구, 서울특별시 구로구, 서울특별시 광진구, 서울특별시 영등포구, 서울특별시 성북구	충청북도 충주시, 전라북도 전주시 덕진구, 경기도 하남시, 경기도 이천시, 충청남도 논산시, 세종특별자치시

2018년 대비 크게 달라진 점이 존재하지 않는다, 2018년과 2021년 모두 수도권, 도시 지역이 순위를 차지했다. 광역시와 수도권 등 대도시 지역 위주로 시중은행이 입지 하는 경향이 여전히 지속되고 있음을 확인할 수 있다.

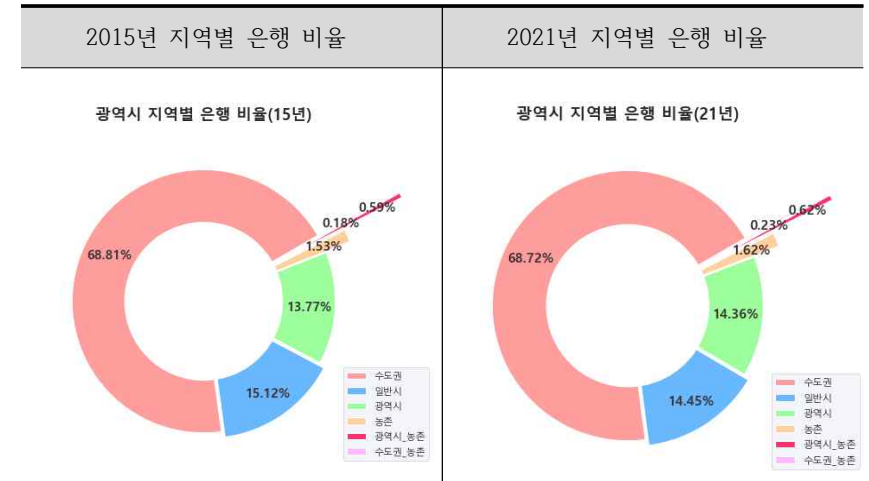
□ 2015 ~ 2018년 시중은행 점포 증감 지도



- 2015, 2021 두 기간 내 시중은행 점포가 없는 군 지역: 46곳
- 전국 250개 지자체 내 46곳에 4대 시중은행이 존재하지 않음
- 분석 대상 시군구: 249곳(경상북도 울릉군 제외)

시중은행이 적게 입지하거나 입지하지 않은 농촌의 경우에는 시중은행 점포 변화량이 없는 경우가 존재하거나 변화량이 크지 않다. 이를 시각화를 통해 쉽게 확인할 수 있다. 일반시의 경우, 신도시와 새로 개발이 되는 도시를 위주로 점포가 증가하는 경우도 존재하나 점포가 증가하는 지역보다는 축소하는 지역이 더 많은 것으로 보인다. 대도시와 같이 인구가 많은 지역의 경우에는 기존에 존재하던 점포를 축소하고 있다, 시각화를 통해 은행이 많이 입지하고 있는 대도시 지역 위주로 은행 점포 감소세가 뚜렷하게 보이는 것을 한눈에 확인할 수 있다.

□ 2015 ~ 2018년 시중은행 점포 증감 시각화



수도권	서울, 경기, 인천
일반시	수도권과 광역시를 제외한 일반 시
광역시	광역시 지역
농촌	군 지역
광역시_농촌	광역시에 속하는 군 지역
수도권_농촌	경기도와 인천에 속하는 군 지역

- 2015 ~ 2021년 내 4대 시중은행 점포의 감소가 일어난 지역 : 은행 점포가 이미 많이 입지 한 광역시도, 수도권과 같은 대도시 지역
- 전국 은행 점포 현황에서 수도권이 차지하는 비중은 60%가 넘는다.
- 도시 이외의 지역에서도 시중은행의 점포 축소 현상은 지속적으로 발생하고 있다.

도시와 농촌의 금융접근성 차이를 알아보기 위해 전국 은행 수를 수도권, 광역시, 일반시, 농촌, 광역시-농촌, 수도권-농촌 이라는 기준으로 나누어 지역별 시중은행 입지 비율을 알아보려고 하였다.

EDA 결과, 수도권이 전국 4대 은행 중 차지하는 비율이 약 68%로 2015년과 2018년 모두 비슷한 수치의 결과를 보였다. 다음으로는 광역시가 약 13%로 두 번째로 높은 비율을 차지했다. 농촌의 경우 약 1% 대로 수도권과 비교했을 때 확연한 차이가 난다. 전국 은행 중 농촌의 4대 시중은행과 수도권 시중은행의 비율에는 큰 차이가 난다. 도시가 차지하는 비율이 높아 이를 같은 특성으로 묶어서 분석했을 시 도시에 관한 특성만 도출될 우려가 있다. 이러한 문제로 인해 도시와 농촌의 금융소외 해소방안을 다르게 보아야 한다.

따라서 비슷한 금융취약지수를 가지고 있는 지역이라도 도시 규모에 따라서 차이가 있을 것이라고 예상되며, 본 분석에서는 ‘도시와 농촌의 금융 취약 특징이 다르다’ 라고 가정하고 분석을 진행한다.

3.2.3 금융취약지수 정의 및 산출

□ 금융취약지수 정의

본 연구에서는 금융취약지수를 지리적 접근성에 한정한다. 따라서 금융취약지수를 ‘지리적 접근성’에 기반하여 산출하는 방식을 택했다. 금융취약지수를 기준으로 금융취약지수에 따른 해결방안을 제시한다.

- 금융취약계층: 고령자, 장애인, 농어민
- 금융접근성: 본 연구에서는 지리적 접근성으로 의미를 한정
- 금융취약성: 지리적 접근성 결여에 따라 금융서비스 이용이 제한되거나 취약한 것을 뜻함. 금융취약지수가 높을수록 금융취약성이 크다.

□ 금융취약지수 산출

금융취약지수 = 시군구 내 그리드별 최단 OD-Matrix 값의 평균

금융취약지수는 500mX500m 그리드로 나누어진 시군구의 중심점을 기준으로 4대 시중은행이 각 그리드의 중심점으로부터 얼마나 멀리 떨어져 있는지에 대한 지수이다. 금융취약지수는 거리를 기준으로 산출된다. 금융취약지수가 높을수록 시중은행이 멀리 위치한다. 본 분석에서는 비대면 금융으로 인해 발생하는 지리적 금융소외를 해결하기 위해 이동이 불편한 노약자와 장애인, 농어민들을 기준으로 금융취약계층을 정의하고, 금융소외를 해소하는 방안중 하나로 금융취약지수 산출을 통해 문제를 해결하고자 한다.

지리적 접근성을 기반으로 금융취약지수를 산출하기 위해 OD Matrix를 계산한다. OD Matrix는 Origin-Destination의 약어로 기종점을 뜻한다. QGIS의 QNEAT3 플러그인을 이용하여 기종점 사이의 최단 경로를 계산한 뒤, 이를 평균내 금융취약지수를 계산한다. 아래의 표는 OD Matrix 결과값 표로 계산한 OD Matrix의 Total_cost를 금융취약지수 산출에 이용한다.

□ OD Matrix 결과값 표

Origin_id	출발지 ID(각 그리드의 중심점)	Destination_id	도착지 ID(각 은행 점포)
Entry_cost	출발지 좌표 → 도로 이동거리	Network_cost	도로 → 도로 이동거리
Exit_cost	도로 → 도착지 좌표 이동거리	Total_cost	총 이동거리

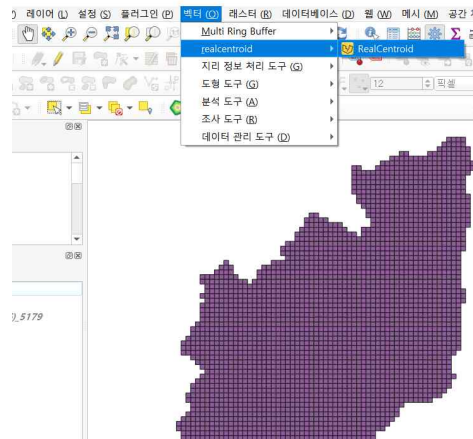
- 총 이동거리(Total_cost)를 금융취약지수 산출에 사용한다.

Step 1. 각 그리드별로 Origin_id(그리드의 중심점)와 Destination_id(은행 점포) 간 총 이동거리(Total cost)의 계산값들 중 최솟값을 산출한다.

Step 2. 각 시군구의 최종 금융취약지수 값은 각 그리드별 최솟값을 평균낸 값으로 설정한다.

□ QGIS 금융취약지수 산출 과정

① 각 행정구역 내 500m x 500m 그리드 & 중심점 생성



② 각 그리드의 중심점(Origin)에서 전국 시중은행 점포(Destination)까지의 도로 기준 최단 거리 계산



③ 계산된 OD Matrix를 바탕으로 Python을 이용하여 각 그리드별 최솟값의 평균 계산

```
[6]: # origin_id(그리드)별 total_cost의 최소값 출력
data = pd.DataFrame(data_od.groupby('origin_id')['total_cost'].min())

[7]: data

[7]:      total_cost
origin_id
77      7444.365389
78      6979.966028
79      7432.336629
80      9234.327842
81      9709.143524
...
11193    9526.787984
11194    8991.135624
11315    6484.446519
11316    11999.798816

[8]: # OD matrix 값 산출 : 그리드 평균
data['total_cost'].mean()

[8]: 9176.811037929037
```

④ 금융취약지수

앞서 산출된 값을 바탕으로 시군구 행정단위별 금융취약지수를 도출

시군구_구분	금융취약지수
인천광역시 옹진군	76.019
전라남도 완도군	58.82
전라남도 신안군	57.26
전라남도 고흥군	56.961
전라남도 진도군	53.29
경상북도 영덕군	50.318
경상북도 영양군	50.219
강원도 평창군	43.947
.....
경기도 수원시 팔달구	0.843
서울특별시 영등포구	0.802
서울특별시 서대문구	0.797
대구광역시 중구	0.738
서울특별시 동대문구	0.658
서울특별시 중구	0.556
울산광역시 중구	0.445

3.2.5 데이터 분석 모델링

□ 분석 대상 선정

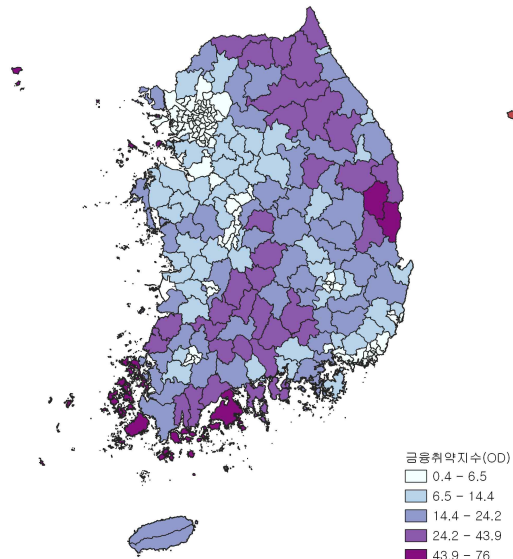
탐색적 분석을 진행한 결과를 바탕으로 세부적인 분석 대상을 선정한다. 지리적 접근성에 한정하여 전국 금융취약지 분석을 진행하며 동시에 농촌과 도시의 금융취약지의 차이에 관해 분석한다.

□ 대상 지역

- 전국 시군구 249개 지역 대상(경상북도 울릉군 제외)
- 도시: 특별시, 광역시, 일반시(시 내의 구 지역도 포함)
- 농촌: 군 지역
- 도시 지역: 168개 / 농촌 지역: 81개

(목적) 지역별 계산된 금융취약지수를 기준으로 도시와 농촌 지역을 나누어서 지역별 특성에 맞는 금융소외 해소방안 마련

□ 전국 금융취약지수 시각화



4. 분석결과

4.1 금융취약지 특성 변수 T-검정

- 독립표본 T-검정 실시하기 위한 가정 탐색

■ 독립성 : 독립변수의 그룹군은 서로 독립적이어야 한다. 도시와 농촌, 도시 내 금융취약/비취약 지역, 농촌 내 금융취약/비취약 지역 간 관계가 없으므로 독립성을 확보하였다.

■ 정규성 : 각 집단의 표본은 정규성을 가져야 한다. 도시와 농촌, 각 집단의 측정값이 20개로 30개 미만이므로 Shapiro-Wilk Test를 진행해 정규성을 검토한다. 이때,

귀무가설(H0) : 표본의 모집단이 정규분포를 따른다.

대립가설(H1) : 표본의 모집단이 정규분포를 따르지 않는다.

■ 등분산성 : 두 집단의 분산은 동일해야 한다. 정규성을 만족하는 경우 Levene's Test를 통해 등분산성 검정을 진행해 등분산성을 만족할 경우 등분산 독립표본 T-검정을, 만족하지 않는 경우 이분산 독립표본 T-검정을 진행한다. 이때,

귀무가설(H0) : 등분산이다.

대립가설(H1) : 등분산이 아니다.

1) 도시-농촌 간 금융취약지수 차이 검정

- 귀무가설(H0) : 도시와 농촌 간 금융취약지수의 평균은 같을 것이다.

대립가설(H1) : 도시와 농촌 간 금융취약지수의 평균은 같지 않을 것이다.

- 정규성을 만족하지 않으므로 맨-휘트니 U 비모수 검정 실시

(출력값) Mann-Whitney U Test Result(statistic = 892.5, pvalue = 1.2324658158910944e-28)

- p-value < 0.05이므로 귀무가설을 기각한다. 따라서 도시와 농촌 간 금융취약지수는 같지 않다.

⇒ 도시와 농촌 사이 금융취약 정도는 분명한 차이가 있으므로 별개로 고려한다.

2) 도시 내 금융 취약지역/비취약지역 간 변수 차이 검정

금융취약지수를 기준으로 도시 상위 20개(취약지역), 하위 20개(비취약지역) 간 외부 특성변수들의 유의미한 차이가 있는지 검정

- 가정 탐색(정규성&등분산성) → Mann-Whitney U Test / 이분산 T-검정 / 등분산 T-검정

귀무가설(H0) : 도시의 금융취약지역과 비취약지역 간 “변수”의 값은 같을 것이다.

대립가설(H1) : 도시의 금융취약지역과 비취약지역 간 “변수”의 값은 같지 않을 것이다.

변수명	정규성	등분산성	검정방법	T-검정/비모수 검정
(전자상거래 이용금액)	0.096/0.007		비모수	0.00
(배달앱 이용금액)	0.48/0.008		비모수	0.00
자가 거주자 수	0.788/0.021		비모수	0.002
주택담보대출 잔액	0.043/0.028		비모수	0.00
신용대출 잔액	0.024/0.007		비모수	0.00
평일 이동거리	0.901/0.022		비모수	0.007
주말 이동거리	0.885/0.02		비모수	0.003
주중 대중교통	0.211/0.00		비모수	0.00
주말 대중교통	0.216/0.00		비모수	0.00
주중 하이패스	0.045/0.071		비모수	0.001
주말 하이패스	0.108/0.062	0.011	이분산 T-검정	0.00
투자 정보	0.418/0.012		비모수	0.00
자동차 정보	0.777/0.014		비모수	0.00
하이엔드 소비 정보	0.353/0.019		비모수	0.00
하이엔드 소비대상자 수	0.001/0.032		비모수	0.001
하이엔드 소득대상자 수	0.00/0.003		비모수	0.00
추정 연소득	0.112/0.012		비모수	0.00
연간 카드 소비 금액	0.068/0.018		비모수	0.00
순자산 평가금액	0.002/0.005		비모수	0.00
국산차량 보유자 수	0.925/0.031		비모수	0.285
수입차량 보유자 수	0.001/0.032		비모수	0.00
사업체 수	0.009/0.003		비모수	0.00
종사자 수	0.001/0.008		비모수	0.001
금융 취약계층 인구수	0.359/0.359	0.077	등분산 T-검정	0.032

(결과)

- 대부분의 특성변수에 대해 귀무가설을 기각, 즉 취약지역과 비취약지역 사이 평균이 같지 않다. 이는 도시 내에서도 취약지역과 비취약지역 사이 분명한 차이가 드러남을 의미한다.

도시 지역 내 금융취약지역과 비취약지역 사이 ‘국산차량 보유자 수’에는 유의미한 차이가 없는 것으로 나타났다. 금융취약지역의 국산차량 보유자 수는 총 4025.1명(평균)/80502명(합), 비취약지역의 경우 총 4897.2명(평균)/97944명(합)이었다.

3) 농촌 내 금융 취약지역/비취약지역 간 변수 차이 검정

금융취약지수를 기준으로 농촌 상위 20개(취약지역), 하위 20개(비취약지역) 간 외부 특성변수들의 유의미한 차이가 있는지 검정

- 가정 탐색(정규성&등분산성) → Mann-Whitney U Test / 이분산 T-검정 / 등분산 T-검정

귀무가설(H0) : 농촌의 금융취약지역과 비취약지역 간 “변수”의 값은 같을 것이다.

대립가설(H1) : 농촌의 금융취약지역과 비취약지역 간 “변수”의 값은 같지 않을 것이다.

변수명	정규성	등분산성	검정방법	T-검정/비모수 검정
(전자상거래 이용금액)	0.072/0.00		비모수	0.068
(배달앱 이용금액)	0.026/0.00		비모수	0.011
자가 거주자 수	0.018/0.00		비모수	0.061
주택담보대출 잔액	0.252/0.00		비모수	9.442e-05
신용대출 잔액	0.052/0.00		비모수	0.002
평일 이동거리	0.336/0.00		비모수	0.01
주말 이동거리	0.534/0.00		비모수	0.011
주중 대중교통	0.026/0.00		비모수	0.007
주말 대중교통	0.028/0.00		비모수	0.007
주중 하이패스	0.002/0.00		비모수	2.420e-05
주말 하이패스	0.049/0.00		비모수	0.000
투자 정보	0.039/0.00		비모수	0.006
자동차 정보	0.049/0.00		비모수	0.048
하이엔드 소비 정보	0.525/0.00		비모수	0.002
하이엔드 소비대상자 수	0.252/0.00		비모수	4.5e-05
하이엔드 소득대상자 수	0.00/0.00		비모수	0.000
추정 연소득	0.083/0.00		비모수	5.969e-05
연간 카드 소비 금액	0.157/0.00		비모수	0.000
순자산 평가금액	0.229/0.00		비모수	4.10e-05
국산차량 보유자 수	0.184/0.00		비모수	0.000
수입차량 보유자 수	0.013/0.00		비모수	6.238e-05
사업체 수	0.295/0.00		비모수	0.000
종사자 수	0.403/0.00		비모수	1.555e-05
금융 취약계층 인구수	0.048/0.0006		비모수	0.000

(결과)

대부분의 특성변수에 대해 귀무가설(H0) 기각, 즉 취약지역과 비취약지역 사이 평균이 같지 않다. 이는 농촌 내에서도 취약지역과 비취약지역 사이 분명한 차이가 드러남을 의미한다.

농촌 지역 내에서도 '전자상거래 이용금액'과 '자가 거주자 수'의 경우 금융취약지역과 비취약지역 사이에 큰 차이가 없는 것으로 확인되었다. 실제 전자상거래 이용금액의 경우 금융취약지역에서는 총 447,740,000,000원, 비취약지역에서는 총 2,173,440,000,000원으로 나타났다. 또한 자가 거주자 수의 경우 금융취약지역에서는 총 462.2명(평균) / 9244명(합), 비취약지역에서는 2181.4명(평균) / 43628명(합)으로 나타났다.

4) 도시와 농촌의 금융취약지역 간 변수 차이 검정

도시와 농촌 각각의 금융취약지역 20개(지수값 상위 20개) 사이 외부 특성변수들의 유의미한 차이 존재하는지 검정

- 가정 탐색(정규성&등분산성) → Mann-Whitney U Test / 이분산 T-검정 / 등분산 T-검정

귀무가설(H0) : 도시와 농촌의 금융취약지역 간 “변수”의 값은 같을 것이다.

대립가설(H1) : 도시와 농촌의 금융취약지역 간 “변수”의 값은 같지 않을 것이다.

변수명	정규성	등분산성	검정방법	T-검정/비모수 검정
(전자상거래 이용금액)	0.007/0.229		비모수	0.00
(배달앱 이용금액)	0.008/0.065		비모수	0.00
자가 거주자 수	0.021/0.018		비모수	0.00
주택담보대출 잔액	0.028/0.253		비모수	0.00
신용대출 잔액	0.007/0.052		비모수	0.00
평일 이동거리	0.022/0.025		비모수	0.00
주말 이동거리	0.02/0.015		비모수	0.00
주중 대중교통	0.000/0.584		비모수	0.00
주말 대중교통	0.000/0.569		비모수	0.00
주중 하이패스	0.071/0.000		비모수	0.00
주말 하이패스	0.062/0.001		비모수	0.00
투자 정보	0.012/0.004		비모수	0.00
자동차 정보	0.014/0.15		비모수	0.00
하이엔드 소비 정보	0.019/0.046		비모수	0.00
하이엔드 소비대상자 수	0.032/0.252		비모수	0.00
하이엔드 소득대상자 수	0.003/0.000		비모수	0.00

추정 연소득	0.012/0.083		비모수	0.00
연간 카드 소비 금액	0.018/0.158		비모수	0.00
순자산 평가금액	0.005/0.229		비모수	0.00
국산차량 보유자 수	0.031/0.185		비모수	0.00
수입차량 보유자 수	0.032/0.013		비모수	0.00
사업체 수	0.003/0.295		비모수	0.00
종사자 수	0.008/0.403		비모수	0.00
금융 취약계층 인구수	0.446/0.018		비모수	0.00

(결과)

모든 특성변수에 대해 비모수검정 결과 귀무가설 기각, 즉 변수별로 도시의 취약지역과 농촌의 취약지역 사이 유의미한 차이가 존재한다. 따라서 도시와 농촌 지역의 금융취약지역에도 분명한 차이가 있다고 볼 수 있다.

⇒ 따라서 도시와 농촌별 금융취약지역의 외부 변수에는 분명한 차이가 존재하며 이는 도시와 농촌 간 환경 특성에 분명한 차이가 존재함을 나타낸다.

추가적으로 전자상거래 이용금액과 배달앱 이용금액 변수의 경우 디지털취약성의 일환으로 고려했으나 오히려 위 변수들이 택배 및 배달과 같은 오프라인 서비스와 밀접한 관련이 있어 지역별 편차가 나타날 가능성이 있으므로 최종 외부 특성 변수에서 제외하였다.

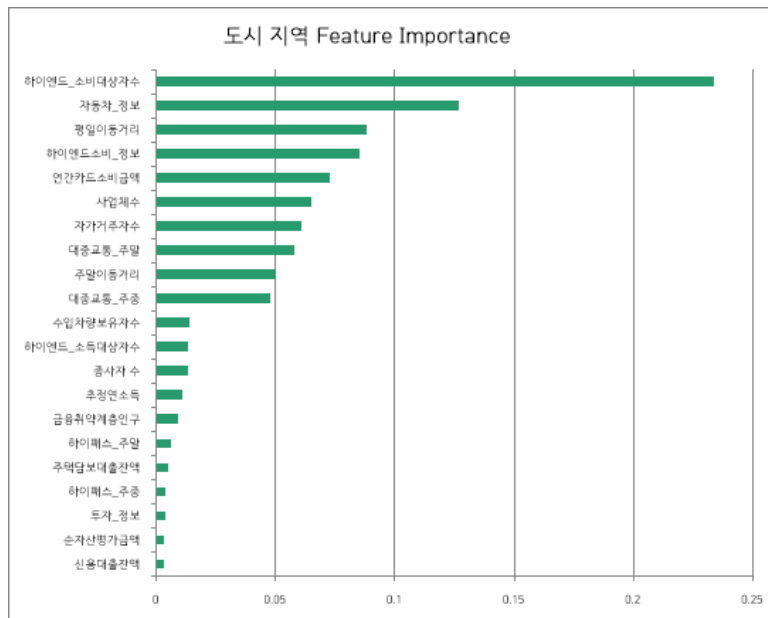
4.2 Random Forest 기반 특성 중요도 산출

- Step 1. 측정 변수들의 단위 차이로 인해 숫자의 스케일이 크게 달라져
Standard -Scaler(평균 0, 분산 1로 조정) 이용해 변수 변환
- Step 2. GridSearch를 통해 금융취약지수 예측 정확도를 높일 수 있는 모델 하
이퍼파라미터 추출
- Step 3. Random Forest 모델을 이용해 도시 금융취약지수 예측
- Step 4. Feature Importance 확인

1) 도시 금융취약지수 특성 중요도 산출

■ 모델 훈련 결과

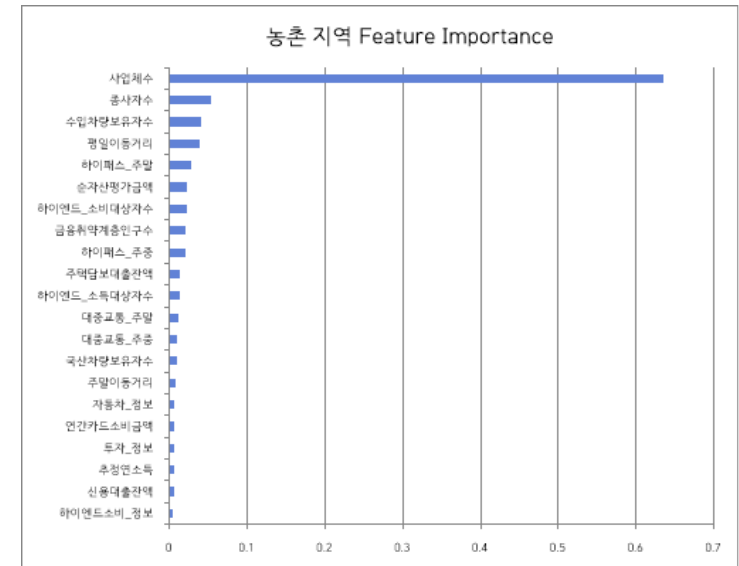
- 최적의 파라미터) max_depth = 6, min_samples_leaf = 3,
min_samples_split = 7, n_estimators = 100
- 예측 정확도 : 70.14%
- 변수별 중요도 결과



2) 농촌 금융취약지수 특성 중요도 산출

■ 모델 훈련 결과

- 최적의 파라미터) max_depth = 6, min_samples_leaf = 3,
min_samples_split = 5, n_estimators = 100
- 예측 정확도 : 67.05%
- 변수별 중요도 결과



4.3 중요도 상위변수 차원축소

변수 간 상관관계 존재하므로 변수 간 차원 축소 진행

- Step 1. 금융취약지수의 평균을 기준으로 금융취약지 선정
- Step 2. Feature Importance 상위 10개 변수를 금융취약지 특성 변수로 선정
- Step 3. Scree Plot/Eigenvalue, 즉 Variance 고려해 축소 차원 개수 선정
- Step 4. PCA 차원 축소

주성분분석(Principal Component Analysis)은 여러 변수 사이 존재하는 상관관계를 이용해 대표하는 주성분을 추출함으로써 정보의 손실을 최소화하는 방향으로 차원을 축소하는 방법론이다. 본 과제에서는 변수 간 다중공선성 문제가 발생 하는 것을 고려하여 주성분 분석을 통해 차원을 축소하고자 한다.

1) 도시 특성 중요도 상위 10개 차원축소

■ 도시의 금융취약지 선정

지역의 금융취약지수가 도시 전체 지역 금융취약지수의 평균보다 작을 경우 “금융취약지”로 판단한다. 결과적으로 총 60개 지역이 선정되었다.

- Feature Importance 결과 기반으로 선정된 변수 10개

하이엔드_소비대상자 수, 자동차_정보, 평일이동거리, 하이엔드소비_정보, 연간카드소비금액, 사업체 수, 자가거주자 수, 대중교통_주말, 주말이동거리, 대중교통_주중

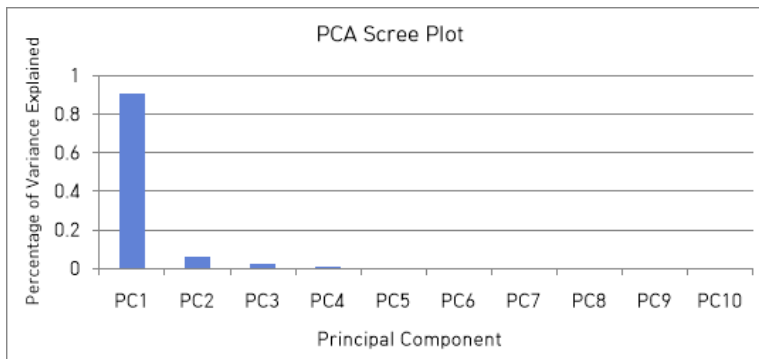
⇒ 도시 지역에서는 소비 및 기동성(이동성, 차량)과 관련된 변수들이 금융취약성에 영향을 주는 것으로 나타났다.

- 축소 차원 개수 선정

기준 1. Eigenvalue(고유값) 0.7 이상 & 설명 가능한 분산(누적기여율) 80% 이상

	설명 분산 비율	기여율	누적기여율
PC1	8.519	0.829	0.829
PC2	0.886	0.086	0.916
PC3	0.410	0.04	0.956
PC4	0.143	0.014	0.97
...			
PC10	0.006	0.001	1.000

기준 2. Scree Plot 이용해 판단!



⇒ 고유값과 Scree Plot을 통해 판단한 결과 축소 차원은 1개로 지정한다.

- 차원 축소 결과

변수	PC1	변수	PC1
하이엔드_소비대상자수	0.327	사업체수	0.293
자동차_정보	0.328	자가거주자수	0.324
평일이동거리	0.328	대중교통_주말	0.289
하이엔드소비_정보	0.322	주말이동거리	0.327
연간카드소비금액	0.330	대중교통_주중	0.290

⇒ 계수의 절대값이 클수록 성분을 구성할 때 해당 변수의 중요성이 크다고 볼 수 있다. 차원 축소 결과 연간카드소비금액, 자동차_정보, 평일이동거리, 하이엔드_소비대상자수, 주말이동거리 변수의 중요성이 많이 반영되었다.

2) 농촌 특성 중요도 상위 10개 차원축소

- 농촌의 금융취약지 선정

지역의 금융취약지수가 농촌 전체 지역 금융취약지수의 평균보다 작을 경우 “금융취약지”로 판단한다. 결과적으로 총 38개 지역이 선정되었다.

- Feature Importance 결과 기반으로 선정된 변수 10개

사업체 수, 종사자 수, 수입차량 보유자 수, 평일이동거리, 하이패스_주말, 순자산평가금액, 하이엔드_소비대상자수, 금융취약계층인구수, 하이패스_주중, 자가거주자수

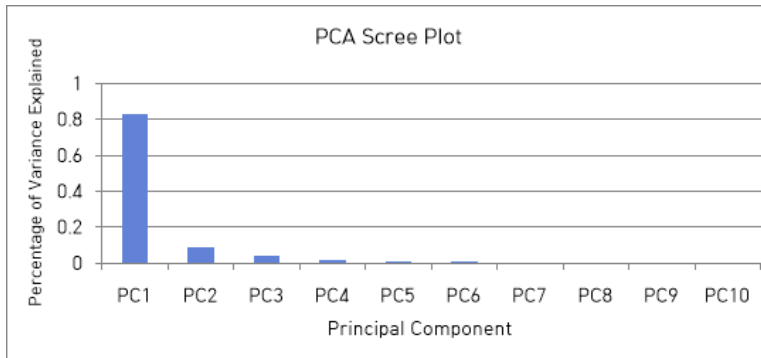
⇒ 은행업 입지 자체에 영향을 주는 사업체, 종사자 변수와 자산규모, 기동성 관련 변수들이 농촌 지역에서 금융취약성과 연관 있는 변수로 선정되었다.

- 축소 차원 개수 선정

기준 1. Eigenvalue(고유값) 0.7 이상 & 설명 가능한 분산(누적기여율) 80% 이상

	설명 분산 비율	기여율	누적기여율
PC1	8.519	0.829	0.829
PC2	0.886	0.086	0.916
PC3	0.410	0.04	0.956
PC4	0.143	0.014	0.97
...			
PC10	0.006	0.001	1.000

기준 2. Scree Plot 이용해 판단!



⇒ 고유값과 Scree Plot을 통해 판단한 결과 축소 차원은 2개로 지정한다.

- 차원 축소 결과

변수	PC1	PC2	변수	PC1	PC2
사업체수	0.315	-0.398	순자산평가금액	0.326	0.253
종사자수	0.327	-0.222	하이엔드_소비대상자수	0.336	-0.008
수입차량보유자수	0.317	0.019	금융취약계층인구수	0.245	-0.637
평일이동거리	0.331	-0.084	하이패스_주중	0.308	0.395
하이패스_주말	0.317	0.348	자가거주자수	0.330	0.194

⇒ 차원 축소 결과 첫 번째 주성분은 모든 변수에 대해 연관성을 가지고 있다. 따라서 이 성분은 해당 지역의 전반적인 금융취약성을 평가하는 주요 지표로 해석할 수 있다. 반면 두 번째 성분은 금융취약계층과 강한 음의 연관성을 보이며 사업체 수와도 음의 연관성을 보이므로 두 번째 성분은 지역의 경제성 활동성을 평가하는 지표로 해석할 수 있다.

4.4 금융취약지 군집화 및 결과도출

Step 1. 실루엣 계수 & Elbow Method를 기준으로 K값 선정

Step 2. K값 기준으로 군집화

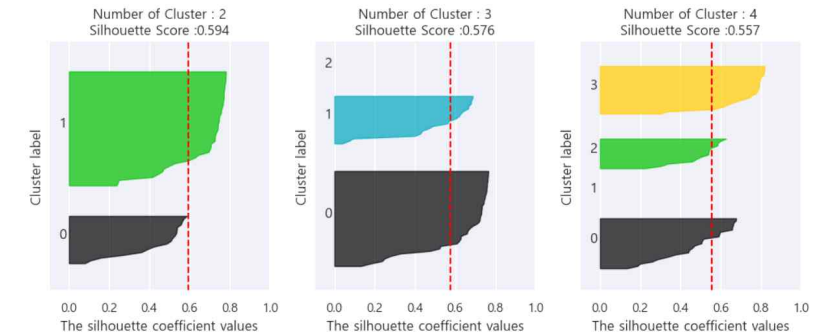
1) 도시 K-Means 기준 군집화 결과

- 실루엣 계수 & Elbow Method를 기준으로 K값 선정

군집화 기준 변수 : PCA를 거친 PC1 변수 & 은행 점포 개수

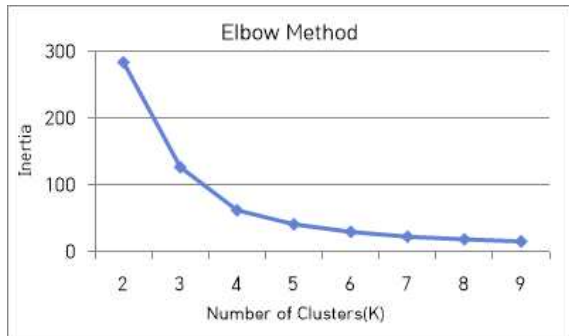
K값 결정 기준 ① 실루엣 지수(Silhouette Index)

Silhouette 지수는 클러스터링 결과를 평가하는 지표로서 각 데이터별로 데이터가 속한 군 내의 유사도와 인접한 군 간의 유사도를 비교하는 지표이다. Silhouette 지수 시각화 결과에서 각 색상은 클러스터를 의미하며 같은 클러스터 내에서 Silhouette 값이 1에 가까운 데이터가 많다면 군집화가 잘 되었다고 볼 수 있다. 또한 만약 클러스터 내 1개의 데이터 포인트만 존재하는 경우, 해당 데이터 포인트의 실루엣 계수는 0으로 본다.



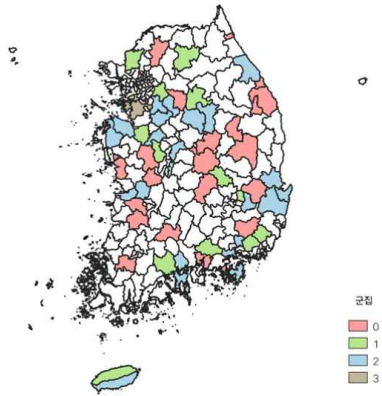
K값 결정 기준 ② Elbow Method(Inertia)

이너셔(inertia)란, 클러스터의 중심과 클러스터에 속한 샘플 사이 거리의 제곱 합으로 클러스터에 속한 샘플이 얼마나 가깝게 모여 있는지를 나타내는 값이다. Elbow Method는 클러스터 개수를 늘려가면서 이너셔의 변화를 관찰해 최적의 클러스터 개수를 찾는 방법이다. 클러스터 개수를 증가시키면 이너셔의 감소 속도가 꺾이는 지점이 있는데 이 지점이 바로 최적의 클러스터 개수이다. 이 지점 이후로는 클러스터 개수를 늘리더라도 클러스터에 밀집된 정도가 크게 개선되지 않으며 오히려 의미 없이 여러 개의 그룹으로 분류되어 모델의 효율이 크게 떨어진다.



⇒ Silhouette 계수와 Elbow Method에 따라 클러스터 개수 K = 4 선택

- K = 4 군집화 결과



군집	도시명
0	강원도 동해시, 강원도 삼척시, 강원도 속초시, 강원도 태백시, 경기도 여주시, 경기도 포천시, 경상남도 밀양시, 경상남도 사천시, 경상남도 통영시, 경상북도 김천시, 경상북도 문경시, 경상북도 상주시, 경상북도 안동시, 경상북도 영주시, 경상북도 영천시, 전라남도 나주시, 전라북도 김제시, 전라북도 남원시, 전라북도 정읍시, 충청남도 공주시, 충청남도 논산시, 충청남도 보령시, 전라남도 순천시, 제주특별자치도 제주시, 강원도 춘천시, 강원도 원주시, 경기도 안산시 단원구, 경상남도 진주시, 경상북도 구미시, 경상남도 양산시, 경기도 파주시, 경기도 광주시, 세종특별자치시, 충청남도 아산시, 경상남도 김해시, 대구광역시 동구
1	전라남도 여수시, 경상북도 포항시 북구, 강원도 강릉시, 제주특별자치도 서귀포시, 충청북도 제천시, 경상북도 경주시, 경상남도 창원시 마산합포구, 충청북도 청주시 상당구, 전라남도 광양시, 경상남도 거제시, 충청북도 충주시, 충청남도 당진시, 경기도 이천시, 충청남도 서산시, 경상북도 포항시 남구, 전라북도 군산시, 경기도 안성시, 경상북도 경산시, 경기도 용인시 처인구, 충청남도 천안시 동남구, 전라북도 익산시, 경상남도 창원시 의창구, 대전광역시 동구
2	전라남도 여수시, 경상북도 포항시 북구, 강원도 강릉시, 제주특별자치도 서귀포시, 충청북도 제천시, 경상북도 경주시, 경상남도 창원시 마산합포구, 충청북도 청주시 상당구, 전라남도 광양시, 경상남도 거제시, 충청북도 충주시, 충청남도 당진시, 경기도 이천시, 충청남도 서산시, 경상북도 포항시 남구, 전라북도 군산시, 경기도 안성시, 경상북도 경산시, 경기도 용인시 처인구, 충청남도 천안시 동남구, 전라북도 익산시, 경상남도 창원시 의창구, 대전광역시 동구
3	경기도 화성시

- 각 변수별 군집 분포 시각화



군집 1 : 소비(하이엔드_소비대상자수, 연간카드소비금액) 규모가 작을 뿐만 아니라 소비재에 대한 욕구가 적고 자가 거주자, 사업체, 주말 이동 또한 적음

군집 2 : 자동차와 고개 상품 등의 재화에 대한 욕구와 관심이 많으며 이에 따라 소비를 많이 하고 자가 거주자가 많으며 주말 이동이 많음

군집 3 : 주중, 주말 내내 대중교통을 많이 이용하며 사업체 중심 도시

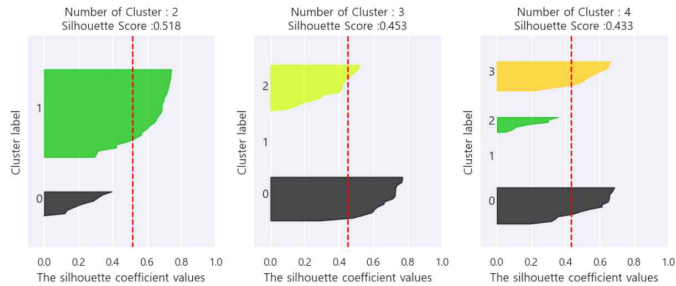
군집 4 : 다른 도시들에 비해 소비, 인구유동성, 사업체 모두 많은 편으로 타 도시들과 차별점을 지님

2) 농촌 K-Means 기준 군집화 결과

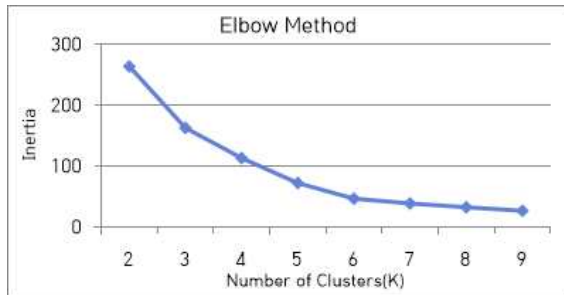
- 실루엣 계수 & Elbow Method를 기준으로 K값 선정

군집화 기준 변수 : PCA를 거친 PC1과 PC2 변수 & 은행 점포 개수

K값 결정 기준 ① 실루엣 지수(Silhouette Index)

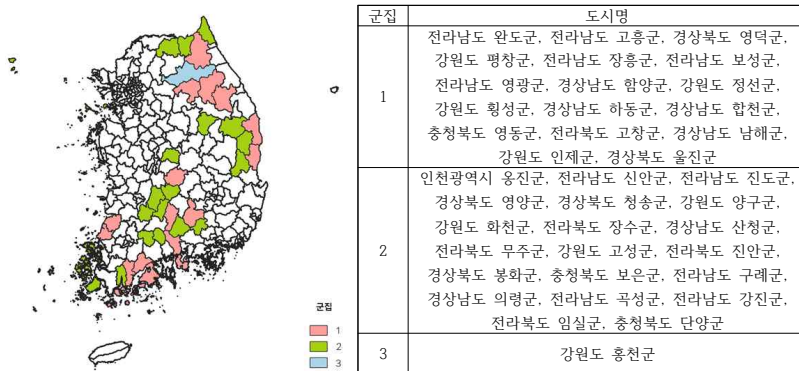


K값 결정 기준 ② Elbow Method

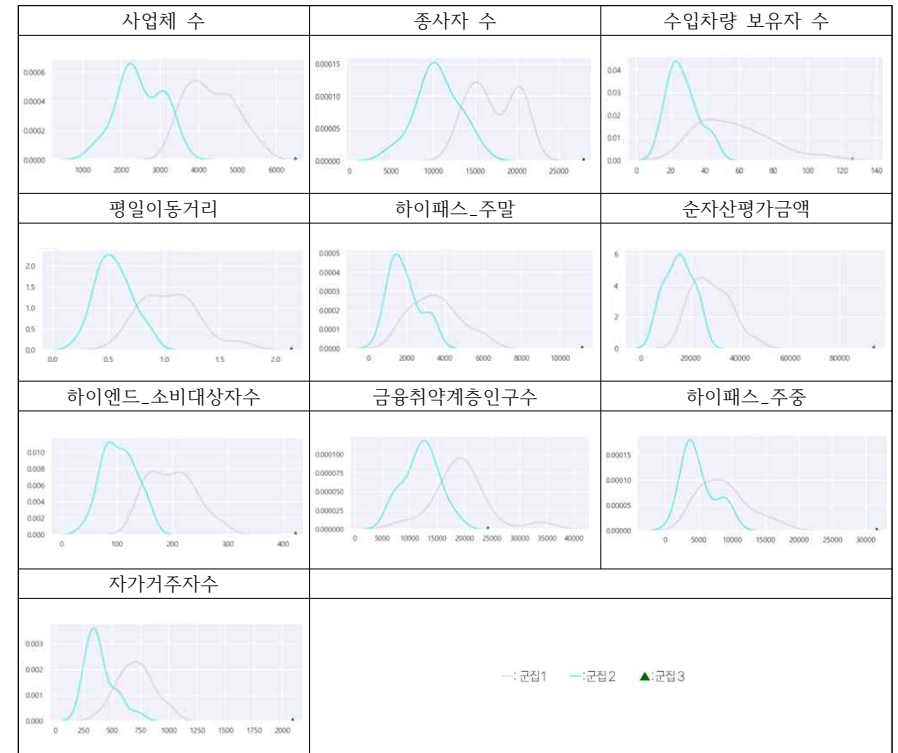


⇒ Silhouette 계수와 Elbow Method에 따라 클러스터 개수 K = 3 선택

- K = 3 군집화 결과



- 각 변수별 군집 분포 시각화



군집 1 : 회사와 공장 등이 2 ~ 3천 개 정도 분포해 있으며(사업체 수), 따라서 종사자 또한 많고 면적이 좁으며 차차 통근이 많은 농촌, 특히 대부분의 사람들이 수입차량을 보유하지 않음

군집 2 : 사업체가 많고 농촌 내 종사자 수도 많은 지역이며 금융취약계층 인구수가 많고 같은 군집 내에서도 고른 분포를 보임, 농촌 내에서도 군집 1이 군집 2에 비해 금융 취약 심화

군집 3 : 다른 도시에 비해 상대적으로 금융취약계층 인구수가 많은 편은 아니나 사업체 및 종사자 수가 많고 차량을 이용한 이동이 많아 다수의 유동 인구 대비 상대적으로 은행 점포 수가 적어 금융취약지로 분류

5. 금융소의 해소방안 및 기대효과

5.1. 금융소의 해소방안

지역별 금융소의 분석 결과 도시 및 농촌은 각각 차별되는 요인에 따라 거리적 금융소의현상이 나타난다. 따라서 정책효과 제고를 위해서 해소방안 도출 시 도시와 농촌으로 나누어 지역적 차별성을 고려한 접근이 필요하다.

도시지역의 금융소의 해소방안은 첫째 지하철 역사 내 은행방문 예약 시스템 설치, 생활 필수재 중심의 금융혜택 활성화, 근로소득자의 비금융 혜택 확대 등이 있다.

첫째, 출퇴근 또는 이동 중에 이용 가능한 역내 방문예약 시스템을 구축하여 본인 근무지 또는 접근이 용이한 곳의 영업점에 미리 방문 시간을 지정하여 접근성을 제고하는 방안 있다. 주로 사업체 또는 종사자 수가 밀집된 근무지를 중심으로 금융기관의 물리적 접근성이 양호하기 때문에 실질적인 수요 및 업무 처리시간 단축을 위해 이동 중에 예약을 하고 지정된 시간에 이용하는 방식이다.

둘째, 도시 지역의 금융취약 현상은 지리적 금융소외보다 개인적 금융소외, 즉 상품과 서비스의 용이성에 영향을 받는 것으로 예상된다. 따라서 결제액에 따른 포인트 부과, 실물지원 등 물가 상승기에 가계경제에 직접적인 혜택을 제공하여 생활수준을 보완하는 형태로 금융소외를 해소할 수 있다.

셋째, 영세 법인 근로자 대상으로 금융·비금융 지원을 확대할 필요가 있다. 근무지에서 은퇴설계, 자산관리 등 비금융 서비스를 연계하여 평균 소득 하위 법인 대상으로 적극적인 소득 보전 및 자산형성 지원을 실시해야 한다.

도시 지역의 거리적 금융소외 해소를 위해서는 유동인구가 많은 대중교통 거점을 활용하여 금융거래 소요시간을 감축할 수 있는 방안이 필요하다. 아울러 거리적 금융소외보다는 개인적 금융소외의 편차가 더 크기 때문에 정책을 적용할 계층을 특정하는 것이 중요하다. 영세 법인에 종사하는 저소득 근로자를 대상으로 직접적인 금융지원이 필요하며 생활비를 보전하여 삶의 질을 균질하게 영위할 수 있도록 소비바우처 지급 등 생활밀착형 금융지원이 필요하다.

농촌지역의 해소방안의 경우 이동형 금융서비스의 활성화, 관공서 및 주민센터 기능 확대, 지역기반 사업장 연계한 지원정책 실시 등이 있다.

첫째, 현재 일부 지방은행을 중심으로 운영 중인 “찾아가는 이동점포”의 확대 운영이 필요하다. 농촌지역의 경우 거리적 접근성이 떨어지기 때문에 노인 복지시설 및 주민센터와 연계를 통해 효과를 극대화할 수 있다. 주거래 금융기관 및 인구밀집도 등을 고려하여 최적 운영주기 및 경로를 탐색하여 운영한다면 금융 접근성의 커버리지를 효과적으로 제고할 수 있다.

둘째, 우체국·주민센터 등 관공서를 중심으로 업무위탁을 활성화하여 대면채널을 운

영한다. 특히 우체국은 국영 금융기관이기 때문에 금융업무를 수행할 수 있는 물적·인적 인프라가 행정기관에 비해 이미 갖춰져 있으므로 시중은행과의 업무위탁 및 제휴를 통해 취급 상품 및 서비스의 범위를 확대하여 제공할 수 있다. 따라서 우체국 관서 네트워크를 활용하여 읍·면 거주 농어민 대상 금융상품 및 서비스의 제공을 원활하게 도모할 수 있다.

셋째, 도시지역과 마찬가지로 영세 사업장 종사자를 대상으로 적극적인 자산관리를 지원하기 위한 은행권 금융 상담 서비스를 연계하여 제공할 수 있다. 농어촌 지역의 사업장을 기준으로 산발적으로 분포되어 있는 인구가 밀집하기 때문에 사업장을 거점으로 이동점포 운영 또는 금융기관과의 제휴 서비스 발굴한다면 금융소의 해소 효과를 극대화할 수 있다.

5.2. 기대효과

첫째, 지리적 금융소의 해소를 위한 정책·제도 마련 시 정량적인 금융취약지수를 활용하여 의사결정을 지원할 수 있다. 지리데이터와 OD 매트릭스 모델링을 통해 도출한 취약지수를 토대로 금융접근성을 판단하고 정책 도입 이후 정량적인 효과 또한 산출 가능하다.

둘째, 도·농 지역별 금융취약성에 영향을 미치는 주요변수를 확인하였다. 지리적 금융소외의 판단에 있어 전국을 대상으로 일원화하는 것이 아니라 판단기준을 다각화하여 접근함으로써 진단에 따른 차별화된 해소방안 수립이 가능하다.

셋째, 우체국과의 업무 연계를 통한 접근성 개선 효과가 크다. 실제로 강원도 평창군의 경우 기존의 OD 매트릭스 도출값은 43.9km였으나 우체국에서 금융업무 위탁을 통한 관서 인프라를 연계할 경우 값이 9.2km로 대폭 감소하여 물리적 접근성 제고효과가 뚜렷하다. 시중은행의 점포폐쇄 지역을 대상으로 읍·면 단위까지 분포되어 있는 우체국 관서를 활용하여 금융거래를 제공할 경우 농어촌 지역을 중심으로 거리적 금융소외를 해결할 수 있을 것으로 본다.

다만 향후 심층연구 수행 시 보완점은 아래와 같다.

금융취약지수 산출 시 거리적 접근성 이외에도 다른 변수를 고려하여 종합적인 취약지수 산출을 고려할 필요가 있다. 또한 OD 매트릭스 값 도출 시, 그리드의 단위를 더 작게 설정하여 지역별로 면밀하게 산출할 필요가 있으며 최종 산출된 클러스터별 입지 분석을 세분화하여 금융취약 원인에 대한 추가 분석의 필요성이 있다.

금융소의 해소는 거리적·개인적 접근성 제고를 위한 정부 주도의 정책지원이 필요한 공적 영역이다. 지리 데이터, 소득 데이터 등 데이터 기반의 정책 의사결정을 통해 정량효과를 산출할 수 있는 방안 수립·운영을 도모한다면 적극적이고 실증적인 금융소외 해소방안을 통해 정책효과를 극대화할 수 있을 것이다.

6. 참고문헌

- 구찬림, 정준호, 방경식 (2015), 은행점포의 경영성과에 영향을 미치는 요인, *부동산학보*, 61(0), 94-107.
- 금융위원회. 은행 오프라인 금융서비스 접근성 제고 방안. 2022.06. <https://www.fsc.go.kr/po010101/77931?srchCtgr=&curPage=&srchKey=&srchText=&srchBeginDt=&srchEndDt>
- 박희연, 김혜린, 김영란. 노년층 금융거래 불편함을 돈으로 환산한다면. *KOSTAT 통계플러스 2022년 봄호(2022)*. 20-31.
- 신경희. 금융소비자의 디지털 소외 현황 및 각국의 대응. *자본시장연구원 자본시장포커스 제 2021-07호*(2021.04.05.)
- 오영환. (2021.07.28.) 디지털 시대를 살아가는 시니어의 금융소외 문제점과 해결방안[주제발표]. 2021 올드(Y-OLD) 이노베이션 포럼[온라인 세미나]. <https://www.youtube.com/watch?v=fWbrSrjdWwY>
- 이구형. 은행권의 점포 축소와 금융소외계층 보호를 위한 과제. *국회입법조사처 이슈와 논점 제 1893호*(2021.11.22.)
- 이연지. 영국, 은행 공유점포 “뱅크허브” 성공사례. *KIRI 리포트 제 541호* (2022.03.07.), 23-25.
- 이정우, 송영규, 전정배(2021). 은행 지점 통·폐합에 따른 전국 금융 접근성 산출 및 취약지 분석 연구. 제 5회 금융보안원 논문공모전, 1-46.
- 이정훈. “디지털 몰라서, 은행 찾느라... ‘2만4600원’ 더 쓰는 노인들” 한겨레. 2022.03.30. https://www.hani.co.kr/arti/economy/economy_general/1036806.html.
- 최희진. “KB국민은행, 고령층 찾아가는 ‘시니어 라운지’ 운영” 경향신문. 2022.07.19. <https://www.khan.co.kr/economy/finance/article/202207191724011#c2b>.
- 홍준호, 오민지, 조용빈, 이경희, 조완섭(2020). 다차원 데이터의 군집분석을 위한 차원축소 방법: 주성분 분석 및 요인분석 비교. *한국빅데이터학회 학회지*, 5(2), 135-143.