



빅데이터를 활용한 소상공인 노후준비에 대한 사회적 인식 연구

홍희정

국민의 주거 행복을 책임지는 지속가능 주택금융의 선도기관

분야

빅데이터를 활용한 소상공인 노후준비에 대한 사회적 인식 연구

2024년 12월



- 저 자 : 홍희정 (051-663-8165 / flyhongs@hf.go.kr)
- 본고의 내용은 필자의 개인 의견으로 한국주택금융공사의 공식적인 견해와 다를 수 있습니다.

목 차

I. 서론	1
II. 이론적 고찰	3
1. 빅데이터	3
2. 네트워크 중심성	5
3. 텍스트마이닝	7
4. 의미연결망 분석	8
5. 선행연구	8
III. 연구설계	10
1. 연구범위	10
2. 자료수집 및 정제	11
3. 분석방법 및 통계패키지	12
IV. 연구분석 및 결과	12
1. 코로나 19 이전 (1기)	12
2. 코로나 19 이후 (2기)	20
3. 코로나 19 이전과 이후 비교 (1기 vs 2기)	27
V. 결론: 연구결과 요약 및 시사점	29
1. 연구결과 요약	29
2. 시사점	31
참고문헌	34

□ 서론

- 코로나 19 이후 사회·경제적으로 많은 변화가 있었고, 특히 장기화된 경기 침체로 소상공인의 경제적·심리적 어려움이 매우 큰 것으로 나타남
- 코로나 19 이후 폐업과 은퇴에 몰린 소상공인의 증가는 향후 복지재정 부담 가중 등 새로운 사회문제로 자리매김 할 수 있으므로, 이에 대한 선제적인 준비가 필요할 것으로 예상
- 코로나 19 이후 소상공인을 지원하기 위한 다양한 정책상품이 쏟아지고 있지만, 이는 대체로 단기적인 방안으로 미래사회를 위한 장기적인 관점에서 지속가능한 지원책을 고민할 필요가 있음
- 따라서 이 연구에서는 소상공인의 노후준비에 대한 사회적 인식을 조사한 후 향후 소상공인의 노후준비를 위한 정책적 시사점을 도출하고자 함

□ 이론적 고찰

- 빅데이터 개념 및 특징
 - (개념) 인터넷에서의 많은 양의 정형화된 데이터와 비정형 데이터를 처리하는 것을 말하며(Sehgal & Agarwal, 2018), 전통적인 데이터 관리 방식으로 저장하거나 분석 혹은 시각화하기 어려운 대량의 다양하고 복잡한 구조로 생성된 데이터 집합을 의미(Sagiroglu & Sinanc, 2013; Coombes & Barry, 2014)
 - (특징) 빅데이터의 특징은 크게 3V로 구분할 수 있음
 - 규모(Volume), 다양성(Variety), 속도(Velocity)
- 네트워크 중심성
 - (개념) 네트워크에 있는 각 노드가 네트워크의 중심에서 얼마나 가까운지를 나타내는 척도로 각 노드의 네트워크 기능에서의 중요도를 나타냄

- (네트워크 척도) 연결정도중심성, 매개중심성, 근접중심성, 위세중심성 등이 있음

- 텍스트마이닝

- 텍스트마이닝은 자연처리 기술을 기반으로 대규모의 텍스트 데이터로부터 의미 있는 정보와 지식을 추출하는 정보기술

- 의미연결망 분석

- 사람 또는 조직의 사회적 관계를 시각화하고 분석하여 개체 간 연결된 정도를 나타내거나 행위자에 의해 인식되는 사회적 관계를 파악하는데 활용

□ 연구설계

- 연구범위

- (시간적) 코로나 19 전후 대비를 위해 1기와 2기로 구분
 - 1기 (코로나 19 이전): 2016.1.1.-2019.12.31.)
 - 2기 (코로나 19 이후): 2020.1.1.-2023.12.31.)
- (내용적) 연구의 명확성을 위해 핵심 키워드로 ‘소상공인 노후’, ‘소상공인 노후준비’로 수집하여 일치 시킴

- 자료수집 및 정제

- (수집채널) 네이버, 다음, 구글 등에서 웹문서, 블로그, 뉴스, 카페, 지식인, 페이스북 등에서 자료수집
- (데이터정제) 텍스트를 통해 전체 데이터 정제 후 불필요한 조사와 어미, 대명사 삭제, 키워드와 관련 없거나 무의미한 단어 삭제 또는 분석에서 배제, 관련성이 높거나 유사의미를 지닌 단어 한 단어로 일치하여 통일

- 분석방법 및 통계패키지

- (분석방법) 텍스트마이닝을 통한 주요 키워드 분석, 네트워크 분석, 의미연결망 분석 실시
- (통계패키지) 텍스트 6.0v, Ucinet 6.0 활용

□ 연구분석 및 결과

● 주요 키워드 분석

- 코로나 19 이전에는 소상공인의 노후준비와 관련된 키워드가 핵심 키워드로 나타났으나, 코로나 19 이후에는 소상공인 지원 관련 키워드가 핵심 키워드로 나타나 코로나 19를 기점으로 소상공인의 노후준비에 대한 인식 변화가 있었음을 유추 할 수 있음

● 네트워크 분석

- 코로나 19 이전보다 코로나 19 이후 소상공인의 경제 상황이 더욱 악화된 것으로 확인됨에 따라 코로나 19를 기점으로 소상공인의 노후 준비 경향이 달라졌음을 유추할 수 있음
 - 1기에서는 없었으나, 2기에서 생계, 위기 등이 네트워크 연결성 및 영향력이 높은 키워드로 등장

● 의미연결망 분석 (구조적 등위성 분석)

- 코로나 19 이전에는 소상공인들이 개인 사업을 하면서 노후준비를 스스로 동시 고려하는 것으로 나타났으나, 코로나 19 이후 노후준비를 포함하여 많은 부분을 소상공인들이 정부지원에 의존하는 것으로 확인
 - 1기 클러스터와 2기 클러스터를 비교하면 2기에 코로나 지원과 정부지원이 동시에 등장함으로 소상공인에 대한 지원범위가 보다 확대되고 적극적으로 변화했음을 유추할 수 있음
 - 1기 클러스터: ① 현 상황에 대한 인식, ② 현 사업에 대한 인식, ③ 노후준비에 대한 인식, ④ 정부지원에 대한 인식
 - 2기 클러스터: ① 노후준비에 대한 인식, ② 코로나 지원에 대한 인식, ③ 현 상황에 대한 인식, ④ 정부지원에 대한 인식

□ 결론: 연구결과 요약 및 시사점

- (연구결과 요약) 코로나 19 이전과 코로나 19 이후 소상공인의 전반적인 상황이 매우 달라졌고, 이로 인해 노후준비에 대한 인식에도 변화가 있었던 것으로 확인

- 코로나 19 이전 소상공인들은 개인의 여력에 따라 노후준비를 하되, 현업과 노후준비를 동시 고려하는 경향이 높았던 것으로 유추할 수 있으나, 코로나 19 이후 어려워진 경제 상황으로 현업과 노후준비를 동시 고려하는 것은 사실상 붕괴된 것으로 유추 됨
- 코로나 19 이후 소상공인의 다수가 노후준비 보다는 현 생계를 위해 정부 지원에 의존하는 경향이 높아짐으로써 소상공인의 향후 노후준비도 국가차원에서 고려해야 할 주요 사안으로 인지되고 있음 유추

● (시사점)

- 소상공인의 노후준비에 대한 정부의 인식 전환 필요
- 소상공인의 노후준비에 대한 대국민적 공익 홍보 강화 필요
- 소상공인의 노후준비를 도울 수 있는 다양한 금융지원책의 보완 필요
- 장기적인 관점에서 소상공인의 지속가능한 지원책으로 주택연금 연계 필요

I. 서론

□ 코로나 19 이후 사회·경제적으로 많은 변화가 있었고, 특히 장기화된 경기침체로 소상공인의 경제적·심리적 어려움이 매우 큰 것으로 나타남

- 코로나 19가 사회의 각계각층에 미친 영향에 대한 연구가 학계를 중심으로 집중 수행되었고, 특히 취약계층에 치명적이었다는 연구결과와 신문보도가 끊이지 않음
 - 2024년 12월 현재 한국학술지인용색인(KCI)에 수록된 코로나 19 이후 사회 변화 및 여러 계층의 변화를 다룬 연구는 약 11,598건으로 나타남¹⁾. 이는 지난 2015년 한국사회를 놀라게 했던 메르스(MERS, Middle East Respiratory Syndrome) 관련 연구²⁾와 비교하면 크게 높은 수치임
 - 주요 언론사는 코로나 19 이후 변화된 사회·경제·문화·정치 등에 대해 연일 보도하고 있으며, 특집기사로 경기침체로 인한 소상공인의 어려움을 집중 보도하기도 함
- 고물가·고금리로 전국민의 가계부채가 증가했고, 특히 장기화된 경기침체로 소상공인 및 자영업자들의 대출이 큰 폭으로 증가했으며, 한계 상황에 놓인 소상공인 및 자영업자들은 노후까지 위협받고 있는 것으로 나타남
 - 중소기업중앙회(2021) 자료에 따르면, 코로나 19 이후 소상공인은 이전보다 정부지원정책의 활용 경험이 증가한 것으로 나타났고, 중소벤처기업부(2024)의 자료에 의하면 소상공인이 가입한 노란우산공제금에서 폐업공제금과 건수는 해마다 증가하여 현재 소상공인의 경영난이 매우 심각한 상황임을 유추할 수 있음. 가령, 2017년 3,700억원 지급건수 5만 2천 건 이었으나, 2023년 공제금은 1조 2,600억 원, 가입건수는 11만 건으로 나타남³⁾

1) 우수등재 568건, KCI 등재 10,633건, KCI 후보 497건 (2024년 12월 검색 기준) www.kci.go.kr

2) 우수등재 2건, KCI 등재 276건, KCI 후보 24건 (2024년 12월 검색 기준) www.kci.go.kr

3) 노란우산공제금은 소상공인에게 주는 퇴직금의 성격으로 대출연체나 국세 체납시에도 압류되지 않음. 그럼에도 불구하고 공제금 해지율은 해마다 증가함으로써 소상공인들의 경영 부담이 매우 크다는 것을 유추할 수 있음. 노란우산 공제액과 지급건수는 코로나 19이전인 2017년 3,700억원(5만 2천 건), 2018년 5,500억원(7만 2천 건), 2019년 6,100억원(7만 5천 건)이었고, 코로나 19이후인 2020년 7,300억원(8만 2천 건), 2021년 9,000억원(9만 5천 건), 2022년 9,700억원(9만 1천 건), 2023년 1조 2,600억원(11만 건)으로 나타남

- 코로나 19 이후 폐업과 은퇴에 몰린 소상공인의 증가는 향후 복지재정 부담 가중 등 새로운 사회문제로 자리매김 할 수 있으므로, 이에 대한 선제적인 준비가 필요할 것으로 예상

 - 코로나 19 이후 어려운 경영난으로 다수의 소상공인들은 가게 및 기업대출을 갚기 위해 노후 자금으로 준비했던 자산이나 예금을 처분하는 경향이 높은 것으로 나타남⁴⁾.

- 코로나 19 이후 소상공인을 지원하기 위한 다양한 정책상품이 쏟아지고 있지만, 이는 대체로 단기적인 방안으로 미래사회를 위한 장기적인 관점에서 지속가능한 지원책을 고민할 필요가 있음

 - 현재 소상공인을 위한 햇살론대환대출 등 다양한 정책상품이 존재하나, 이는 현재 직면한 어려움에 대처하는 단기적인 방안이며, 장기적인 관점에서 소상공인의 안정적인 노후생활을 위한 상품은 현재 미미함
 - 중소기업중앙회(2019)에서 실시한 설문조사에 의하면, 소상공인들은 코로나 19 이후 노후준비에 대한 필요성은 이전보다 더욱 크게 느끼지만, 노후생활 안정을 위한 준비는 여력이 되지 않는다고 응답한 비율이 매우 높은 것으로 나타남

- 따라서 이 연구에서는 소상공인의 노후준비에 대한 사회적 인식을 조사한 후 향후 소상공인의 노후준비에 대한 정책적 시사점을 도출하고자 함

 - 특히 이 연구에서는 소상공인의 노후준비로 최후의 보루인 주택을 활용하여 생계와 주거가 모두 해결될 수 있는 최선책으로 주택연금과의 연계를 고려함으로써 정책적 시사점을 도출하고자 함

4) 국제신문 2024.5.19. “임대료, 빚에 허덕여...‘환갑의 사장님’들 노후자금 깎는다, kookje.co.kr/news2011/asp/newsbody.asp?key=20240520.22012005235

II. 이론적 고찰

1. 빅데이터(Big Data)

□ 개념

- 빅데이터 분석이란 인터넷에서의 많은 양의 정형화된 데이터와 비정형 데이터를 처리하는 것을 말하며(Sehgal & Agarwal, 2018), 전통적인 데이터 관리 방식으로 저장하거나 분석 혹은 시각화하기 어려운 대량의 다양하고 복잡한 구조로 생성된 대량의 데이터 집합을 의미하기도 함(Sagiroglu & Sinanc, 2013; Coombes & Barry, 2014).
- 전통 데이터와 빅데이터의 구분은 <표 1>과 같음

<표 1> 전통 데이터와 빅데이터의 구분

구분	전통적 데이터	빅데이터
수집	<ul style="list-style-type: none"> 통계청과 같은 데이터관리 기관 	<ul style="list-style-type: none"> 다양한 경로를 통한 수집
분석가능 표본	<ul style="list-style-type: none"> 모집단에서 표본 추출 	<ul style="list-style-type: none"> 모집단 전체
데이터 표본	<ul style="list-style-type: none"> 데이터만으로 기본 정보 파악 가능 목적에 따른 정제된 데이터 	<ul style="list-style-type: none"> 불순물이 많아 데이터만으로 기본 정보 파악 불가능
데이터 형식	<ul style="list-style-type: none"> 구조화/정형화 고정적 데이터 수치로 객관화한 형태 	<ul style="list-style-type: none"> 비구조화/비정형화 실시간 생성 및 확장 가능 데이터 역동성, 데이터 간 상호의존성, 관련성 텍스트, 숫자, 동영상 등 다양한 형태
데이터 분류	<ul style="list-style-type: none"> 규칙적인 관계적 데이터베이스로 분류 가능 	<ul style="list-style-type: none"> 규칙적인 관계적 데이터로 분류 불가능
데이터 분석	<ul style="list-style-type: none"> 계량적 모형기반 전통적인 분석기법 통계적 분석 	<ul style="list-style-type: none"> 기계학습 등을 통한 다양한 기법 통계적 분석 데이터 시각화 자연어 처리 텍스트 분석

장점	<ul style="list-style-type: none"> 통제된 정확성 	<ul style="list-style-type: none"> 인공지능 분석 실시간 공개로 인한 투명성 확보
사용의 한계점	<ul style="list-style-type: none"> 제한된 표본 자료수집 비율 	<ul style="list-style-type: none"> 불순물 신뢰성 결여 개인정보와 사생활 침해 국가와 조직의 보안 문제 정보의 불균형 심화 가능 일반화 가능성에 대한 지속적 논의

※ 출처: 김선영(2020) 재인용

- 빅데이터에서 ‘빅(Big)’은 중요성, 복잡성, 도전성을 내포하고 있지만 양적으로 치우쳐 생각되는 경향(Ward & Barker, 2013)이 있어 빅데이터 개념을 정립하고자 하는 학자들 사이에서 ‘라지(Large)’ 데이터와의 구분을 강조함
- 라지데이터와 빅데이터와의 구분은 <표 2>와 같음

<표 2> 라지데이터와 빅데이터 구분

구분	라지데이터(Large Data)	빅데이터(Big Data)
원천	Internal	Internal & External
형태	Nearly structured	Mostly unstructured
취합	Well organized	Nearly unorganized
분석방법	Modeling & Optimization centered	Visualization & Exploratory analysis centered
분석환경	DW + Server	Distributed process + Cloud

출처: 강용희(2021), Textom과 함께하는 빅데이터 아카데미 강의자료에서 발췌

□ 특징

- 빅데이터의 특징은 크게 3V(Volume, Variety, Velocity)로 구분할 수 있음
 - 데이터가 과거에 비해 더 빨리 늘어나고(규모, Volume), 기존의 구조화된 정형데이터 및 비정형데이터를 포함하여 다양한 형태를 지니고 있으며(다양성, Variety), 실시간에 가까운 속도(Velocity) 생성되어 유통되며, 그 속도 또한 매우 빠름(문지영·김학선·이종호, 2020)

2. 네트워크 중심성 (Network Centrality)

□ 개념

- 네트워크 중심성이란 네트워크에 있는 각 노드가 네트워크의 중심에서 얼마나 가까운지를 나타내는 척도. 각 노드의 네트워크 기능에서의 중요도를 나타냄
 - 네트워크 중심성은 소셜 네트워크 분석(social network analysis)에서 처음 등장했으며, 이후 네트워크 개념이 접목되는 다양한 분야로 확장되어 네트워크의 기능에서 중요한 노드가 무엇인지 밝혀내기 위한 목적으로 사용 됨

□ 척도

- 네트워크 중심성 척도는 대표적으로 연결정도 중심성(degree centrality), 매개 중심성(betweenness centrality), 근접 중심성(closeness centrality), 위세 중심성(eigenvector centrality)이 있음(Freeman, 1979; Koschutzki & Schreiber, 2008)
- 연결정도 중심성
 - 연결정도 중심성이란 가장 단순하지만 신뢰성 있는 개념으로 각 노드에 대해 그 노드가 상호 작용하는 이웃 노드의 개수(또는 에지의 개수)로 정의. 따라서 이웃 노드가 많을수록 연결정도 중심성이 높아짐으로 직접적인 영향력 측정 가능
 - 연결정도 중심성이 큰 노드들은 네트워크에서 이들이 없어진다면 에지로 연결되어 있는 이웃 노드들과 상호작용이 없어져 기능에 손실을 야기하기도 함
 - 에지 별로 가중치가 존재하는 네트워크의 경우, 모든 에지의 가중치의 합으로 정의되기도 하며, 에지의 방향성이 존재하는 네트워크의 경우 노드로부터 나가는 에지(네트워크 기능에 더 영향을 미침)와 노드를 향해 들어오는 에지의 개수로 구분하기도 함
 - 연결정도 중심성은 전체 네트워크를 고려하지 않고, 이웃한 노드들과의 관계만을 고려하는 개념이기 때문에 전체 네트워크의 중심으로부터 멀리 떨어져 있더라도 연결된 에지가 많으면 중요하다고 판단될 수 있는 한계가 있음

● 매개 중심성

- 네트워크 상의 두 개의 노드 사이의 최단 경로를 이용하는 개념. 네트워크 상의 두 개의 다른 노드 세트들(set 1, 2)을 정의하고, set 1의 노드와 set 2의 노드의 모든 조합에 대한 최단 경로들 상에 각 노드가 얼마나 자주 등장하는지를 정량화 함
- 이는 어떤 노드가 set 1, 2의 최단 경로 상에 많이 존재한다면, set 1, 2사이의 관계를 정의할 때 중요한 역할을 할 것이라는 가정에서 나온 개념. 매개 중심성은 특정 set 1, 2의 상호작용에 있어서 중요한 역할을 할 것이라는 가정에서 나온 개념으로 특정 set 1, 2의 상호작용에 있어서 중요한 노드를 동정할 때 사용
- 연결정도 중심성과 같이 네트워크 전체 기능에서 중요한 노드를 결정할 때도 사용할 수 있으며 노드 x 에 대해 x 를 제외한 나머지 모든 노드들 각 쌍에 대해 (s, t) 모든 최단 경로를 구하고, 그 중 노드 x 를 거치는 최단 경로의 비율을 계산한 뒤, 이를 모든 노드 쌍들에 대해 더해주는 아래의 정의를 따름

$$C(x) = \sum_{s \neq x \neq t} \frac{\sigma_{st}(x)}{\sigma_{st}}$$

- 여기서 σ_{st} 는 노드 s 와 노드 t 사이의 최단 경로의 총 개수를 의미하고, $\sigma_{st}(x)$ 는 그 중 x 를 거치는 최단 경로의 개수를 의미

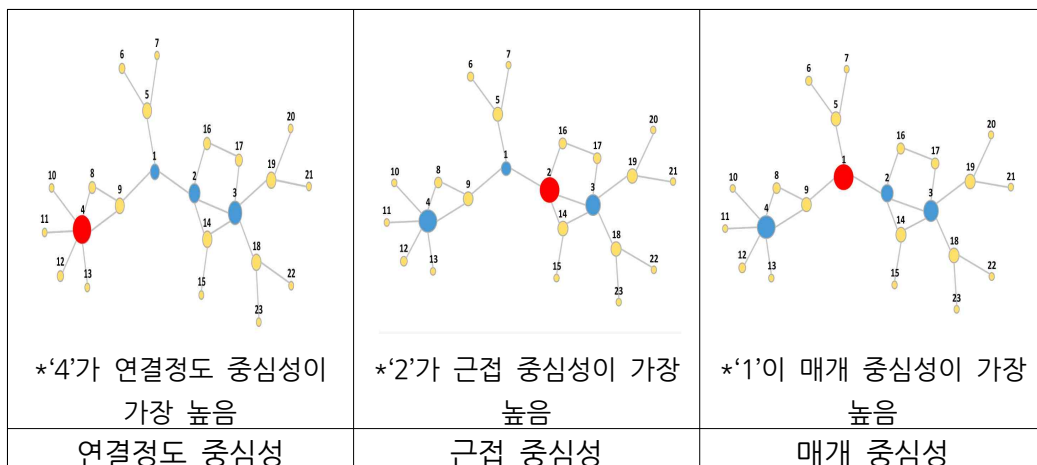
● 근접 중심성

- 한 노드가 다른 노드들까지 도달하는 경로가 짧을수록 중요한 노드일 것이라는 가정에서 시작된 개념. 네트워크상의 노드 세트(set)를 결정하고, 각 노드에 대해서 세트 내의 모든 노드들과의 최단 경로를 계산한 후, 그들의 평균값의 역수로 근접 중심성이 정의
- 이 개념이 확장되어 연결정도 중심성과 같이 네트워크 전체의 기능에서 중요한 노드를 선정할 때도 사용할 수 있음. 노드 x 가 있을 때 x 로부터 x 를 제외한 다른 모든 노드들까지의 최단 경로의 평균을 구하고, 이에 역수를 취하는 아래의 정의를 따름

$$C(x) = \frac{N-1}{\sum_{y \neq x} d(y, x)}$$

- 여기서 $d(y, x)$ 는 노드 x 와 y 사이의 거리를 의미하며, N 은 전체 네트워크 상의 노드의 개수를 의미
- 위세 중심성
 - 다른 노드의 중심성을 반영하여 중요한 노드와의 연결이 많을수록 중요한 노드로 가정하는 것을 의미
- 네트워크 중심성에 대한 간략한 이해는 <그림 1>과 같음

<그림 1> 네트워크 중심성



※ 출처: 강용희(2021), Textom과 함께하는 빅데이터 아카데미 강의자료

3. 텍스트마이닝(Text Mining)

- 텍스트마이닝은 자연처리 기술을 기반으로 대규모의 텍스트 데이터로부터 의미 있는 정보와 지식을 추출하는 정보기술 임. 대용량의 데이터에서 사용자가 관심을 가지는 정보를 키워드의 수준이 아니라 컨텍스트(Context) 수준의 의미를 찾아내는 프로세스(Fan et al, 2014).
- 텍스트마이닝은 사람에 의해 작성된 텍스트데이터와 비텍스트적인 컨텍스트를 기반으로 언어, 내용, 작성자, 다른 변수의 추론 등의 영역으로 구성

되며, 텍스트 분석 방법에는 분류, 군집, 정보추출, 네트워크 분석, 감성분석, 토픽 분석 등이 포함(Bhargavi et al, 2008; Fan et al., 2014).

- 텍스트마이닝 과정은 총 4단계로 이루어짐
 - ① 데이터수집, ② 데이터전처리(단어를 추출할 수 있는 데이터 소스를 가공), ③ 단어추출(수학적 모델이나 알고리즘을 활용하여 필요한 단계를 추출), ④ 정보 분석(최종단어에서 우선순위 도출, 요약, 군집화)의 순서로 이루어짐

4. 의미연결망 분석

- 의미연결망 분석은 사람 또는 조직의 사회적 관계를 시각화하고 분석하여 개체 간 연결된 정도를 나타내거나 행위자에 의해 인식되는 사회적 관계를 파악하는데 활용(Racherla & Hu, 2009; 최윤정·권상희, 2014)
- 텍스트를 이용한 의미연결망 분석은 전체문서에서 빈도수가 높은 단어의 상대적 영향력의 크기를 파악하여 단어들 간의 관계에서 특정 단어가 어떤 역할을 하는지 분석하여 의미를 찾아내는 것이 중요
- 군집분석: 키워드의 군집화를 진행하여 생성된 단어 그룹 내 주요 주제를 도출하는 방법으로 핵심은 관련성이 높은 단어들의 연관된 속성을 찾는 것이며, 맥락적 고려를 할 수 있음

5. 선행연구

- 이 연구에서는 텍스트마이닝 기법을 활용한 사회적 인식 조사 관련 연구와 소상공인 노후준비에 대한 연구를 구분하여 살펴봄
- 한국연구재단 한국학술지인용색인(KCI)에 등재된 논문을 중심으로 살펴보면 <표 3>과 같음⁵⁾

5) 내용이 방대하므로 2015년 이후 수록된 논 보고서 주제와 관련된 논문만 선별 수록

〈표 3〉 텍스트마이닝 및 소상공인 노후 관련 선행연구

구분	저자	논문제목	연구목적
텍스트 마이닝 활용 (인식조사)	이지백 (2024)	텍스트마이닝을 활용한 부산 부동산 시장 인식 변화 연구	텍스트마이닝 기법을 활용하여 부산 부동산 시장에 대한 대중의 인식 변화를 분석
	성미영 외 3인 (2021)	텍스트 마이닝 기법을 활용한 아동 재난안전정보에 대한 사회적 인식 분석	우리나라에서 발생하는 지진과 같은 대규모 재난 발생 시 아동기 자녀를 둔 가정에서 필요로 하는 재난안전정보 유형을 탐색해보고 기초자료를 제공하는 것
	이재무·박인수 (2021)	텍스트마이닝을 활용한 사회적기업에 대한 사회적 인식 분석	사회적기업을 대상으로 관련 뉴스 기사와 각종 소셜네트워크서비스(SNS) 상 의견을 수집하여 현실적인 사회적 인식 수준을 파악하고자 수행
	이용규 (2022)	SNS빅데이터를 활용한 코로나-19 팬데믹 발생 전, 후의 여행 마케팅 타겟의 트렌드 변화 연구	SNS빅데이터를 이용하여 코로나-19 팬데믹 발생 이전과 이후의 기간에 여행마케팅 타겟의 트렌드 변화를 조사, 분석하고 여행과 관련된 정부의 정책수립이나 입법기관의 입법 시는 물론 여행업계의 여행 관련 마케팅 전략 수립에 유용한 정보를 제공
	이광일·남정민 (2024)	텍스트 마이닝 기법을 활용한 코로나-19 팬데믹 발생 이전과 이후 스타트업 성장요인에 관한 인식변화 연구	텍스트마이닝 기법을 활용하여 코로나-19 팬데믹 이전과 이후 기간에 나타난 스타트업 성장요인에 대한 사회적 인식 변화를 조사, 분석하고 코로나-19 이후의 시대에 스타트업 성장을 위한 정부기관의 정책 수립과 기업의 의사결정에 유용한 정보를 제공
홍희정 (2022)	빅데이터 분석을 통한 초장기 정책모기지에 대한 사회적 인식 연구	텍스트마이닝을 기반으로 한 빅데이터 분석을 통해 초장기 정책모기지에 대한 사회적 인식을 살펴보고 정책적 시사점 도출	
소상공인 노후관련	전병욱 (2017)	소기업·소상공인 공제제도의 문제점과 개선방안	소기업·소상공인 공제제도의 근거법령 및 과세 방식 등에 대한 제도적 분석을 바탕으로 가입재상을 확대시킬 수 있는 합리적이고 구체적인 개선방안 제시
	김윤영 (2020)	불안정 노동자의 노후소득보장: 연금유형에 따른 불안정 노동자의 연금소득 시뮬레이션 분석	탈산업 사회에서 구조적으로 양산되는 불안정 노동자에 주목하여 다층연금체계의 유형에 따라 불안정 노동자의 노후소득보장수준이 어떻게 달라지는지 분석
	김수환·김형규 (2017)	소상공인의 사회안전망에 관한 연구	소상공인 및 1인 자영업자를 모두 포괄하는 약 670여만 소상공인의 사회안전망이 확대될 필요성과 그 방안을 제언
	전병욱 (2016)	노후보장 저축제도의 과세제도 개선방안	노후생활 보장에 공헌하는 유사연금제도인 소기업·소상공인공제, 직장공제회 및 저축성 보험의 실질에 보다 부합하도록 과세체계를 개편함으로써 이들 제도에 대한 국민들의 인식을 개선하는 것과 함께 이들 상품으로부터 얻는 이익에 대하여 연금소득 수준의 조세혜택을 제공함으로써 노후보장 재원의 마련

			을 보다 촉진시킬 수 있는 제도 개선방안을 제시
정수원 (2013)	일본 소규모기업의 공제제도 현황과 과제		일본 소규모기업 경영자의 공제제도에 대한 실태에 대해서 알아보고, 현재 우리나라에서 운영하고 있는 소상공인 공제제도의 운영과 발전의 시사점 도출
문현경·김혜진 (2023)	연금 개혁의 주요 행위자로서 사용자에 대한 기초연구		2018-19년 경사노위 연금 특위에 참여한 사용자 협회(한국경총, 대한상의, 소상공인연합회)를 대상으로 연금제도, 보험료율 및 소득대체율 수준 등에 대한 사용자의 인식과 선호를 파악하는 목적을 가짐

※ 저자구성

□ 선행연구와의 차별성

- 기존 설문조사나 신문기사 외 소상공인의 노후준비와 관련하여 빅데이터를 활용하여 대중의 인식을 전반적으로 분석한 학술 연구는 이 연구가 최초임
- 현재 소상공인 노후준비와 관련된 선행연구 자체가 미미한 상황에서 이 연구는 관련 연구의 기초 자료로 활용될 수 있고, 향후 공사 상품과의 연계를 논하는 측면에서 정책적 기여도가 높음

III. 연구설계

1. 연구범위

- 시간적 범위
 - 코로나 19 전후 대비를 비교하기 위해 코로나 발생 전 1기와 발생 후 2기로 구분 (1기: 2016.1.1.-2019.12.31., 2기: 2020.1.1.-2023.12.31.)
- 내용적 범위
 - 연구의 명확성을 위해 ‘소상공인 노후’, ‘소상공인 노후준비’를 핵심 키워드로 자료를 수집했으며, 최종 ‘소상공인 노후준비’로 수집자료 일치

2. 자료수집 및 정제

□ 자료수집

- 수집채널은 네이버, 다음, 구글 등 한국의 대표적인 SNS 포털사이트 임. 해당 포털사이트는 대한민국에서 현재 가장 활발히 대중에게 활용되고 있다는 점에서 대표성을 지닌다고 판단함
 - 네이버(웹문서, 블로그, 뉴스, 카페, 지식IN)
 - 다음(웹문서, 블로그, 뉴스, 카페)
 - 구글(웹문서, 뉴스, 페이스북)

□ 데이터 정제

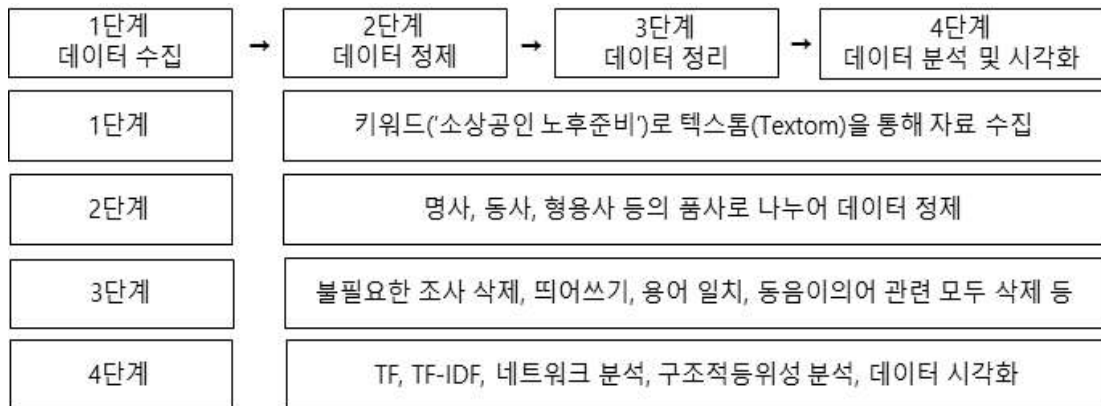
- 텍스트를 통해 전체 데이터 정제를 실시했으며 불필요한 조사와 어미, 대명사를 삭제
 - ‘있’, ‘곳’, ‘등’, ‘것’, ‘그’ ‘등’, ‘ 경우’, ‘이상’ 등 삭제
- 키워드와 관련이 없거나 무의미한 단어는 전문가 의견을 반영하여 소거 또는 분석에서 배제
 - 키워드와 상관없는 동음이의어(同音異議語) ‘노후’와 관련된 모든 내용 삭제
 - 주요 키워드로 단어빈도수를 높게 차지하는 ‘노후’, ‘소상공인’, ‘상공’ 등은 분석에서 배제
- 관련성이 높거나 유사 의미를 지닌 단어는 한 단어로 일치하여 통일
 - ‘노후대비’, ‘노후대책’ 등은 ‘노후준비’로 일치
 - ‘노란우산공제’, ‘노란공제’ 등은 ‘우산공제’로 일치
 - ‘상공인 정책’, ‘상공지원 정책’ 등은 ‘소상공인정책’으로 일치

3. 분석방법 및 통계패키지

□ 분석방법

- 텍스트마이닝, 네트워크 분석, 의미연결망 분석 실시
- 네트워크 분석 흐름도는 <그림 2>와 같음

<그림 2> 네트워크 분석 흐름도



□ 통계패키지

- 데이터 정제 및 네트워크 분석, 감성분석은 텍스트롬 6.0v 활용, 의미연결망 분석은 Ucinet 6.0 활용

IV. 연구분석 및 결과

1. 코로나 19 이전 (1기)

1) 키워드 빈도분석 및 TF-IDF 분석

- 문서 내 추출된 키워드는 총 27,277개이며, 빈도수 상위 10개 키워드는 대출, 노후준비, 창업, 중소기업, 자영업, 정부, 연금, 전통시장, 우산공제, 금리 등임
- 1기의 빈도수 상위 50개 키워드를 시각화하면 <그림 3>과 같으며, 이를 구체적으로 정리하면 <표 4>와 같음

<그림 3> 빈도분석 시 상위 50개 키워드 워드클라우드



<표 4> 주요 키워드 상위 50개

순위	단어	빈도	순위	단어	빈도
1	대출	3052	26	주택담보	382
2	노후준비	2928	27	노후자금	380
3	창업	2017	28	연금저축	372
4	중소기업	1569	29	매출	310
5	자영업	1558	30	보증	307
6	정부	1509	31	노후연금	307
7	연금	1324	32	아파트담보	306
8	전통시장	1200	33	퇴직연금	292
9	우산공제	1191	34	창업자금	283
10	금리	1080	35	개인사업자	282
11	보험	922	36	대환	276
12	소득	890	37	조례	275
13	소기업	825	38	소상공인지원	272
14	국민연금	791	39	사업자대출	244
15	햇살론	632	40	빚	222
16	폐업	627	41	융자	212
17	지원사업	604	42	자자체	211
18	소득공제	595	43	금융지원	209
19	담보	575	44	영세소상공인	204
20	노후생활	572	45	파산	201
21	상공지원	519	46	노후보장	192
22	전세자금	476	47	세액공제	175
23	개인회생	428	48	노후설계	151
24	일자리지원	390	49	소상공인대출	151
25	상환	388	50	전세자금대출	151

□ TF-IDF 분석* 시, 상위 10개 키워드는 대출, 창업, 노후준비, 연금, 자영업, 중소기업, 정부, 우산공제, 전통시장, 금리 등임

● TF-IDF 분석 결과, 상위 50개 키워드를 시각화하면 <그림 4>와 같으며, 이를 구체적으로 정리하면 <표 5>와 같음

* TF-IDF란 'Term Frequency-Inverse Document Frequency'의 약자로 단어 빈도(TF)와 역문서 빈도(IDF)를 곱하여 계산한 것으로 조우제(2022)는 단순히 높은 빈도가 높은 중요도를 의미하지 않고, 문서군 내에서 자주 사용되는 경우, 이는 그 단어가 흔하게 나타나는 것이라 말하며, TF-IDF 분석의 중요성을 강조

<그림 4> TF-IDF 분석 시 상위 100개 키워드 워드클라우드



<표 5> TF-IDF 분석 시 상위 50개

순위	단어	TF-IDF	순위	단어	TF-IDF
1	대출	6487.746	26	일자리지원	1436.596
2	창업	4911.207	27	주택담보	1414.845
3	노후준비	4897.061	28	노후자금	1400.848
4	연금	3889.341	29	조례	1287.468
5	자영업	3807.696	30	노후연금	1245.412
6	중소기업	3759.821	31	매출	1234.058
7	정부	3663.183	32	퇴직연금	1230.382
8	우산공제	3637.11	33	아파트담보	1215.79
9	전통시장	3612.762	34	보증	1212.811
10	금리	3030.191	35	창업자금	1171.29
11	보험	2861.668	36	개인사업자	1131.408
12	소득	2731.558	37	대환	1131.35

13	소기업	2575.252	38	소상공인지원	1084.886
14	국민연금	2573.641	39	사업자대출	1059.696
15	햇살론	2176.253	40	빚	960.5
16	소득공제	2147.514	41	융자	918.39
17	폐업	2130.124	42	금융지원	911.183
18	지원사업	2036.629	43	지자체	907.251
19	담보	1916.431	44	파산	900.262
20	노후생활	1864.289	45	영세소상공인	866.634
21	상공지원	1764.328	46	노후보장	838.15
22	전세자금	1695.811	47	세액공제	822.032
23	개인회생	1608.29	48	전세자금대출	719.121
24	연금저축	1524.863	49	소상공인대출	711.693
25	상환	1443.915	50	노후설계	705.771

□ 빈도분석과 TF-IDF 분석 결과 비교

- 빈도분석에서 높은 비율을 차지한 노후준비, 정부, 전통시장, 폐업, 지원 사업, 일자리지원 등의 키워드는 TF-IDF 분석 시 값이 낮아지는 것으로 확인
- TF-IDF 분석 후 빈도분석과 비교하여 값이 높아진 키워드는 창업, 연금, 우산공제, 노후연금, 퇴직연금 등은 높아지는 것으로 확인
- 빈도분석과 TF-IDF 분석결과를 비교하여 정리하면 <표 6>과 같음

<표 6> 빈도분석과 TF-IDF 분석결과 비교

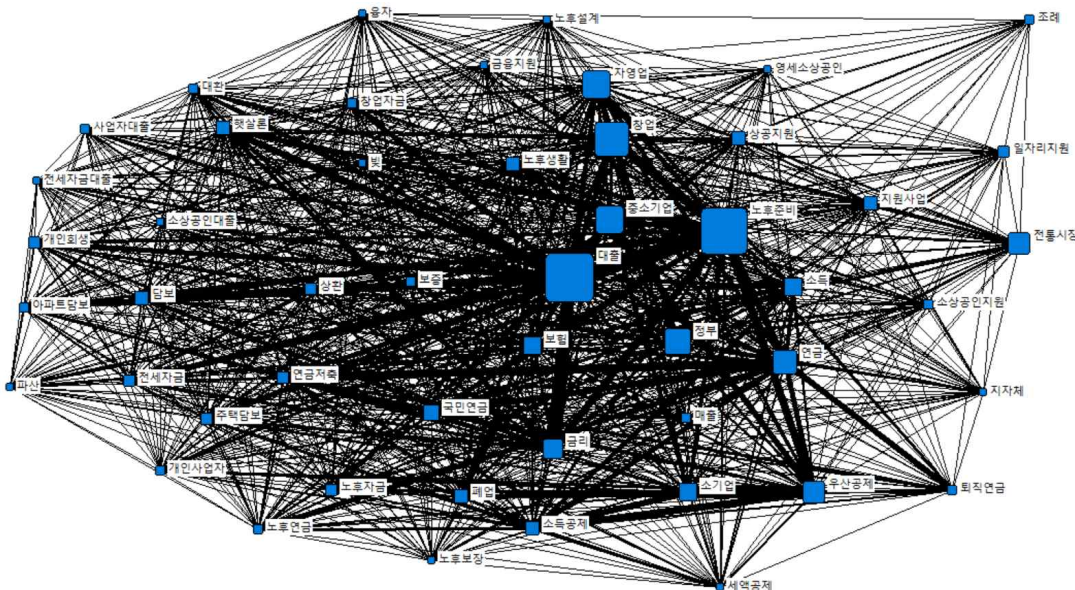
순위변동없음	순위상승	순위하락
대출	창업	노후준비
금리	연금	정부
보험	중소기업	전통시장
소득	우산공제	폐업
소기업	소득공제	지원사업
국민연금	연금저축	일자리지원
햇살론	조례	주택담보
담보	노후연금	노후자금
노후생활	퇴직연금	매출
상공지원	금융지원	아파트담보
전세자금	파산	보증
개인회생	전세자금대출	창업자금
상환		개인사업자
소상공인지원		대환
사업자대출		지자체
빚		영세소상공인
융자		노후설계
노후보장		
세액공제		
소상공인대출		

2) 네트워크 분석 (중심성 분석)

□ 네트워크 상의 주요 키워드를 시각화하면 <그림 5>과 같음

- 노드(node)는 키워드 빈도에 따라 사이즈가 결정되고, 라인(line)은 연결성에 따라 결정
 - 빈도가 높을수록 노드가 크게 표현되고, 연결성이 높을수록 굵은선으로 표현

<그림 5> 1기 네트워크 분석의 주요 키워드 시각화



□ 연관 키워드의 연결중심성, 근접중심성, 매개중심성, 위세중심성 등 네트워크 특성을 분석한 결과는 <표 7>과 같음

- 네트워크 분석결과⁶⁾, 연결중심성이 높은 키워드 상위 10개는 대출, 노후준비, 창업, 중소기업, 자영업, 정부, 연금, 전통시장, 우산공제, 금리 등으로 나타남
- 위세중심성이 높은 키워드 상위 10개는 대출, 노후준비, 창업, 중소기업, 자영업, 정부, 연금, 보험, 소득, 소기업 등으로 나타남

6) 근접중심성과 매개중심성은 값의 차이가 크지 않아 해석 생략

- 한편, 빈도 순위는 낮지만 연결중심성이 높은 키워드로 보험, 소득, 국민연금, 노후생활, 매출, 보증 등이 있으며, 빈도 순위는 높지만 연결중심성이 낮은 키워드로 전통시장, 우산공제, 금리 등이 있음. 빈도순위는 낮지만 위세중심성이 높은 키워드로 보험, 소득, 소기업 등 있으며 빈도순위는 높지만 위세중심성이 낮은 키워드로 보험, 소득, 소기업 등이 있음

〈표 7〉 1기 네트워크 분석

순위	단어	연결중심성	근접 중심성	매개 중심성	위세 중심성
1	대출	0.999	1	0.004	0.155
2	노후준비	0.999	1	0.004	0.155
3	창업	0.999	1	0.004	0.155
4	중소기업	0.999	1	0.004	0.155
5	자영업	0.999	1	0.004	0.155
6	정부	0.999	1	0.004	0.155
7	연금	0.999	1	0.004	0.155
8	전통시장	0.755	0.803	0.001	0.119
9	우산공제	0.816	0.844	0.001	0.129
10	금리	0.979	0.98	0.002	0.153
11	보험	0.999	1	0.004	0.155
12	소득	0.999	1	0.004	0.155
13	소기업	0.918	0.924	0.002	0.155
14	국민연금	0.979	0.98	0.002	0.144
15	햇살론	0.897	0.907	0.001	0.153
16	폐업	0.959	0.960	0.002	0.143
17	지원사업	0.897	0.907	0.002	0.150
18	소득공제	0.938	0.942	0.002	0.141
19	담보	0.897	0.907	0.001	0.147
20	노후생활	0.979	0.98	0.003	0.143
21	상공지원	0.959	0.960	0.003	0.151
22	전세자금	0.897	0.907	0.001	0.149
23	개인회생	0.836	0.859	0.000	0.134
24	일자리지원	0.693	0.765	0.001	0.109
25	상환	0.959	0.960	0.002	0.151
26	주택담보	0.918	0.924	0.001	0.145
27	노후자금	0.918	0.924	0.002	0.144
28	연금저축	0.938	0.942	0.003	0.148
29	매출	0.979	0.98	0.002	0.152
30	보증	0.979	0.98	0.002	0.153
31	노후연금	0.877	0.890	0.002	0.137
32	아파트담보	0.857	0.875	0.001	0.137
33	퇴직연금	0.673	0.758	0.000	0.107
34	창업자금	0.938	0.942	0.002	0.148
35	개인사업자	0.897	0.907	0.001	0.142
36	대환	0.877	0.890	0.001	0.140
37	조례	0.408	0.628	0.000	0.065
38	소상공인지원	0.877	0.890	0.002	0.137
39	사업자대출	0.795	0.830	0.000	0.128

40	빚	0.938	0.942	0.002	0.148
41	융자	0.857	0.875	0.002	0.134
42	지자체	0.775	0.816	0.001	0.123
43	금융지원	0.938	0.942	0.003	0.146
44	영세소상공인	0.836	0.859	0.002	0.131
45	파산	0.795	0.830	0.000	0.128
46	노후보장	0.857	0.875	0.001	0.135
47	세액공제	0.673	0.753	0.000	0.109
48	노후설계	0.857	0.875	0.001	0.136
49	소상공인대출	0.877	0.890	0.001	0.140
50	전세자금대출	0.816	0.844	0.000	0.132

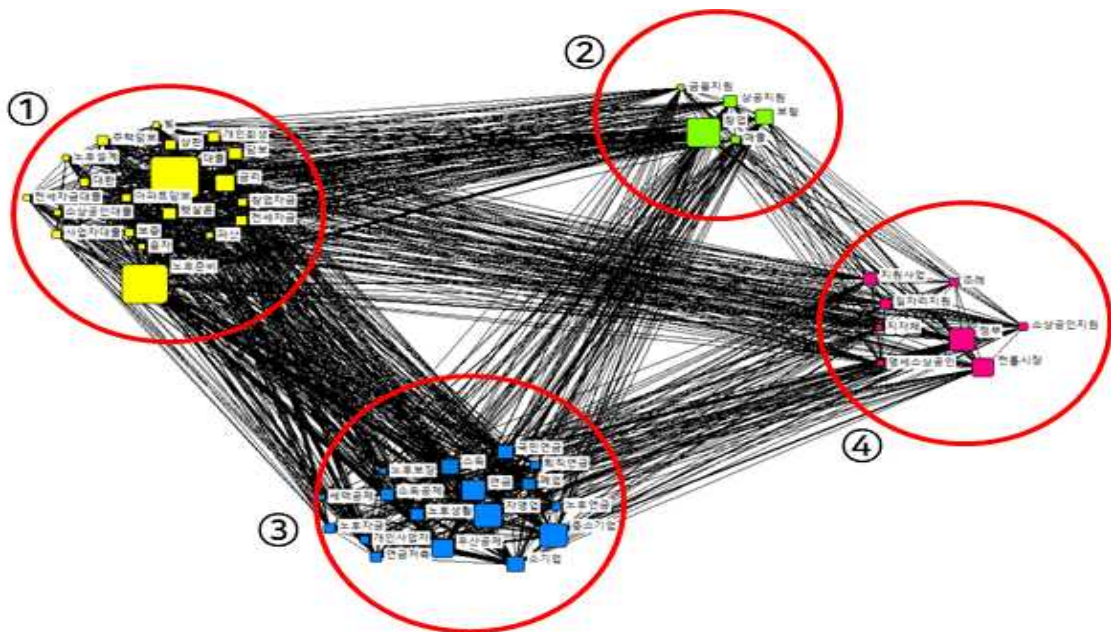
3) 의미연결망 분석

□ 구조적 등위성 분석

- 덴드로그램(Dendrogram)을 통하여 3단계 위계적 집단을 확인
 - 구조적 등위성 분석을 하여 시각화한 결과 <그림 6>과 같고, 키워드 클러스터는 <표 8>과 같음
- 클러스터는 4개가 형성되었고, 각각 다음과 같이 명명함. ① 소상공인의 현 상황에 대한 인식, ② 소상공인의 현 사업에 대한 인식, ③ 소상공인의 노후준비에 대한 인식, ④ 소상공인의 정부지원에 대한 인식
 - 그룹 ①은 대체로 소상공인들이 자신들의 사업을 새롭게 시작하거나 확장 등을 위해 다양한 대출이나 창업 자금에 대해 주로 알아보는 키워드로 구성된 것을 확인, 그룹 ②는 그룹 ①과 유사하나 실질적인 개인 대출 등이 아닌 보다 소극적인 방식의 소상공인 지원 유무를 확인하는 과정으로 확인 됨, 그룹 ③은 소상공인의 노후준비와 관련된 키워드들로 우산공제, 노후자금, 노후생활, 노후연금 등의 키워드가 포함되어 있고, 그룹 ④는 정부지원 관련한 키워드가 주로 포함 된 것으로 확인 됨
- <그림 >에서 보는 바와 같이 네트워크 연관성은 ①↔③간 가장 높은 것으로 나타나며, 일부 키워드의 경우 그룹 간 연관성이 매우 높은 것으로 나타나지만, 대체로 일반적인 연관성의 정도로 나타남
 - 가령, ④ 그룹에 포함된 소상공인지원은 ①, ②, ③ 그룹의 개별 키워드들과 네트워크 연관성이 높은 것으로 확인

- 구조적 등위성 분석을 통해 소상공인의 현 상황에 대한 인식과 노후준비에 대한 인식의 연관성이 높은 것으로 나타나, 코로나 19 이전에는 소상공인들이 노후준비에 대한 관심이 높고, 동시 진행되었을 가능성이 높은 것을 유추할 수 있음

〈그림 6〉 1기 구조적 등위성 분석 결과 시각화



〈표 8〉 1기 구조적 등위성 분석 주요 키워드 클러스터

	클러스터	핵심단어
1	현 상황에 대한 인식	주택담보, 빚, 상환, 대출, 개인회생, 담보, 금리, 창업자금, 전세자금, 파산, 햇살론, 아파트담보, 노후설계, 대환, 전세자금대출, 소상공인대출, 사업자대출, 융자, 보증, 노후준비
2	현 사업에 대한 인식	금융지원, 상공지원, 창업, 매출, 보험
3	노후준비에 대한 인식	세액공제, 노후자금, 개인사업자, 연금저축, 노후보장, 소득공제, 노후생활, 우산공제, 소득, 연금, 자영업, 소기업, 국민연금, 퇴직연금, 폐업, 노후연금, 중소기업
4	정부지원에 대한 인식	지원사업, 조례, 일자리지원, 지자체, 정부, 영세소상공인, 전통시장, 소상공인지원

2. 코로나 19 이후 (2기)

1) 빈도분석 및 TF-IDF 분석

- 문서 내 추출된 키워드는 총 25,711개이며, 빈도수 상위 10개 키워드는 대출, 노후준비, 창업, 중소기업, 자영업, 정부, 연금, 전통시장, 우산공제, 금리 등임
- 빈도수 상위 50개 키워드를 시각화하면 <그림 7>과 같으며, 이를 구체적으로 정리하면 <표 9>와 같음

<그림 7> 빈도분석 시 상위 50개 키워드 워드클라우드



〈표 9〉 주요 키워드 상위 50개

순위	단어	빈도	순위	단어	빈도
1	지원사업	3318	26	위기	556
2	코로나	3273	27	조례	545
3	창업	3058	28	지자체	526
4	정부	2726	29	소기업	525
5	노후준비	2538	30	취약계층	523
6	보조금지원	2300	31	추가지원	521
7	정부지원금	2182	32	소득공제	475
8	자영업	1876	33	보험료	466
9	중소기업	1791	34	임대인	440
10	대출	1428	35	빚	355
11	전통시장	1402	36	감면	345
12	우산공제	1393	37	보증	332
13	차상위계층	1119	38	재난지원금	314
14	보험	1091	39	퇴직연금	278
15	기초생활수급	1061	40	개인회생	277
16	공제	981	41	기초연금	271
17	생계	971	42	개인사업자	261
18	저소득층	922	43	파산	257
19	임대료	904	44	소상공인지원	240
20	폐업	887	45	상황	226
21	연금	855	46	임차인	204
22	국민연금	795	47	노후연금	202
23	상공지원	763	48	융자	199
24	매출	569	49	금융지원	199
25	노후생활	564	50	연금저축	189

- TF-IDF 분석 시, 상위 10개 키워드는 창업, 지원사업, 코로나, 보조금지원, 정부, 노후준비, 정부지원금, 우산공제, 자영업, 중소기업 등임
- TF-IDF 분석 결과, 상위 50개 키워드를 시각화하면 〈그림 8〉과 같으며, 이를 구체적으로 정리하면 〈표 10〉과 같음

〈그림 8〉 TF-IDF 분석 시 상위 50개 키워드 워드클라우드



〈표 10〉 TF-IDF 분석 시 상위 50개

순위	단어	TF-IDF	순위	단어	TF-IDF
1	창업	8166.225	26	노후생활	2043.737
2	지원사업	7154.384	27	위기	2028.255
3	코로나	6459.107	28	소득공제	1986.575
4	보조금지원	5896.685	29	소기업	1975.527
5	정부	5872.871	30	추가지원	1968.928
6	노후준비	5535.309	31	취약계층	1957.291
7	정부지원금	5432.455	32	지자체	1933.019
8	우산공제	4743.604	33	보험료	1917.436
9	자영업	4669.41	34	임대인	1885.53
10	중소기업	4606.679	35	빚	1514.466
11	대출	4501.619	36	감면	1464.023
12	전통시장	4414.416	37	보증	1430.53
13	보험	3767.29	38	재난지원금	1354.22
14	공제	3447.132	39	갱니회생	1331.765
15	임대료	3320.245	40	퇴직연금	1307.283
16	차상위계층	3137.04	41	기초연금	1244.224
17	기초생활수급	3088.25	42	파산	1244.148
18	연금	2990.701	43	개인사업자	1170.95
19	생계	2957.58	44	상환	1057.499
20	폐업	2955.902	45	소상공인지원	1051.904
21	국민연금	2921.578	46	임차인	966.631
22	저소득층	2806.263	47	연금저축	956.176
23	상공지원	2530.232	48	금융지원	939.332
24	조례	2450.277	49	노후연금	937.231
25	매출	2088.65	50	융자	928.887

□ 빈도분석과 TF-IDF 분석 결과 비교

- 빈도분석에서 높은 비율을 차지한 지원사업, 코로나, 정부, 노후준비, 우산 공제, 중소기업 등의 키워드는 TF-IDF 분석 시 값이 낮아지는 것으로 확인
- TF-IDF 분석 후 빈도분석과 비교하여 값이 높아진 키워드는 창업, 보조금 지원, 자영업, 보험, 공제, 임대료, 연금 등은 높아지는 것으로 확인
- 빈도분석과 TF-IDF 분석결과를 비교하여 정리하면 〈표 11〉과 같음

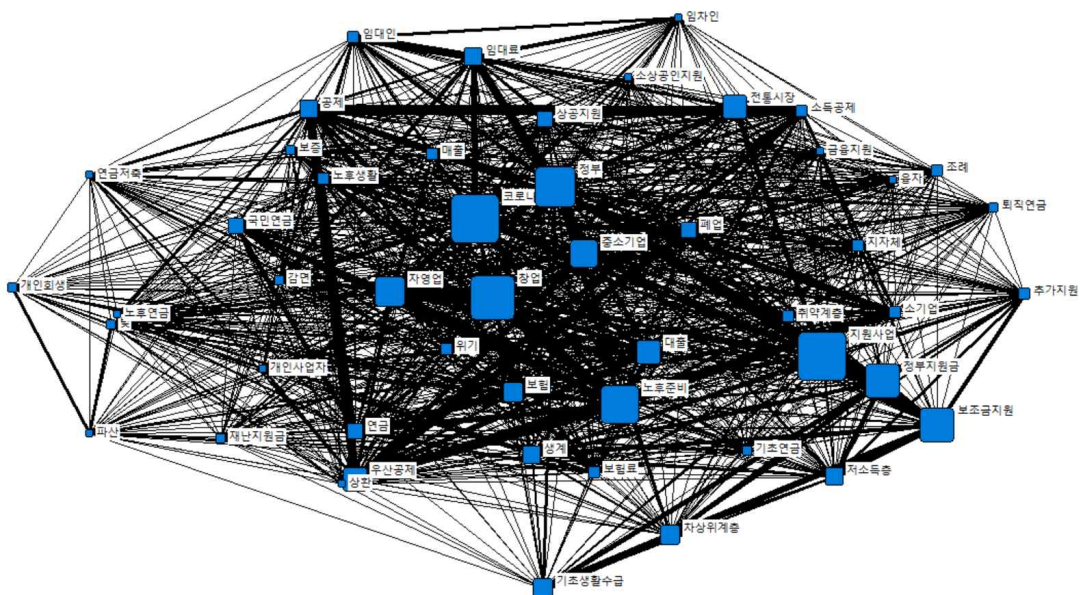
〈표 11〉 빈도분석과 TF-IDF 분석결과 비교

순위변동없음	순위상승	순위하락
정부지원금	창업	지원사업
폐업	보조금지원	코로나
상공지원	자영업	정부
소기업	보험	노후준비
보험료	공제	우산공제
임대인	임대료	중소기업
빚	연금	대출
감면	국민연금	전통시장
보증	조례	차상위계층
재난지원금	소득공제	기초생활수급
기초연금	추가지원	생계
임차인	개인회생	저손그층
	파산	매출
	상환	노후생활
	연금저축	위기
	금융지원	취약계층
		지자체
		퇴직연금
		개인사업자
		소상공인지원
		노후연금
		용자

2) 네트워크 분석 (중심성 분석)

▣ 네트워크 상의 주요 키워드를 시각화하면 <그림 9>와 같음

〈그림 9〉 2기 네트워크 분석의 주요 키워드 시각화



□ **연관 키워드의 연결중심성, 근접중심성, 매개중심성, 위세중심성 등 네트워크 특성을 분석한 결과는 <표 12>와 같음**

- 네트워크 분석결과, 연결중심성이 높은 키워드 상위 10개는 지원사업, 코로나, 창업, 정부, 노후준비, 자영업, 중소기업, 대출 등으로 나타남
- 위세중심성이 높은 키워드 상위 10개는 코로나, 창업, 정부, 자영업, 중소기업, 대출, 생계, 폐업, 위기 등으로 나타남
- 한편, 빈도 순위는 낮지만 연결중심성이 높은 키워드는 보험, 생계, 폐업, 위기 등이 있고, 빈도 순위는 높지만 연결중심성이 낮은 키워드로 지원사업, 보조금지원이 있음. 빈도 순위는 높지만 연결중심성이 낮은 키워드로 지원사업, 보조금지원, 정부지원금이 있음
- 종합적으로 빈도 순위는 다소 낮으나 연결중심성, 근접중심성, 위세중심성이 높은 키워드로 생계, 폐업, 위기 키워드가 확인 됨

<표 12> 2기 네트워크 분석

순위	단어	연결중심성	근접 중심성	매개 중심성	위세 중심성
1	지원사업	0.938	0.942	0.002	0.147
2	코로나	0.999	1	0.003	0.155
3	창업	0.999	1	0.003	0.155
4	정부	0.999	1	0.003	0.155
5	노후준비	0.999	1	0.003	0.155
6	보조금지원	0.836	0.859	0.001	0.133
7	정부지원금	0.938	0.942	0.002	0.147
8	자영업	0.999	1	0.003	0.155
9	중소기업	0.999	1	0.003	0.155
10	대출	0.999	1	0.003	0.155
11	전통시장	0.897	0.907	0.001	0.141
12	우산공제	0.897	0.907	0.002	0.141
13	차상위계층	0.816	0.844	0.001	0.129
14	보험	0.999	1	0.003	0.155
15	기초생활수급	0.591	0.710	0.000	0.094
16	공제	0.938	0.942	0.002	0.147
17	생계	0.999	1	0.003	0.155
18	저소득층	0.877	0.890	0.001	0.137
19	임대료	0.877	0.890	0.001	0.138
20	폐업	0.999	1	0.003	0.155
21	연금	0.938	0.942	0.002	0.147
22	국민연금	0.938	0.942	0.002	0.147
23	상공지원	0.979	0.98	0.002	0.153
24	매출	0.979	0.98	0.002	0.153

25	노후생활	0.979	0.98	0.002	0.153
26	위기	0.999	1	0.003	0.155
27	조례	0.755	0.803	0.000	0.121
28	지자체	0.918	0.924	0.002	0.144
29	소기업	0.897	0.907	0.001	0.141
30	취약계층	0.959	0.960	0.002	0.150
31	추가지원	0.836	0.859	0.001	0.133
32	소득공제	0.857	0.875	0.001	0.136
33	보험료	0.959	0.960	0.002	0.149
34	임대인	0.795	0.830	0.001	0.126
35	빚	0.816	0.844	0.001	0.129
36	감면	0.979	0.98	0.003	0.152
37	보증	0.959	0.960	0.002	0.149
38	재난지원금	0.877	0.890	0.001	0.138
39	퇴직연금	0.816	0.844	0.001	0.129
40	개인회생	0.632	0.731	0.000	0.101
41	기초연금	0.918	0.924	0.002	0.143
42	개인사업자	0.918	0.924	0.002	0.142
43	파산	0.755	0.803	0.001	0.119
44	소상공인지원	0.877	0.890	0.001	0.138
45	상황	0.877	0.890	0.002	0.137
46	임차인	0.714	0.777	0.000	0.115
47	노후연금	0.775	0.816	0.000	0.124
48	융자	0.877	0.890	0.001	0.138
49	금융지원	0.816	0.844	0.000	0.131
50	연금저축	0.755	0.803	0.000	0.120

3) 의미연결망 분석

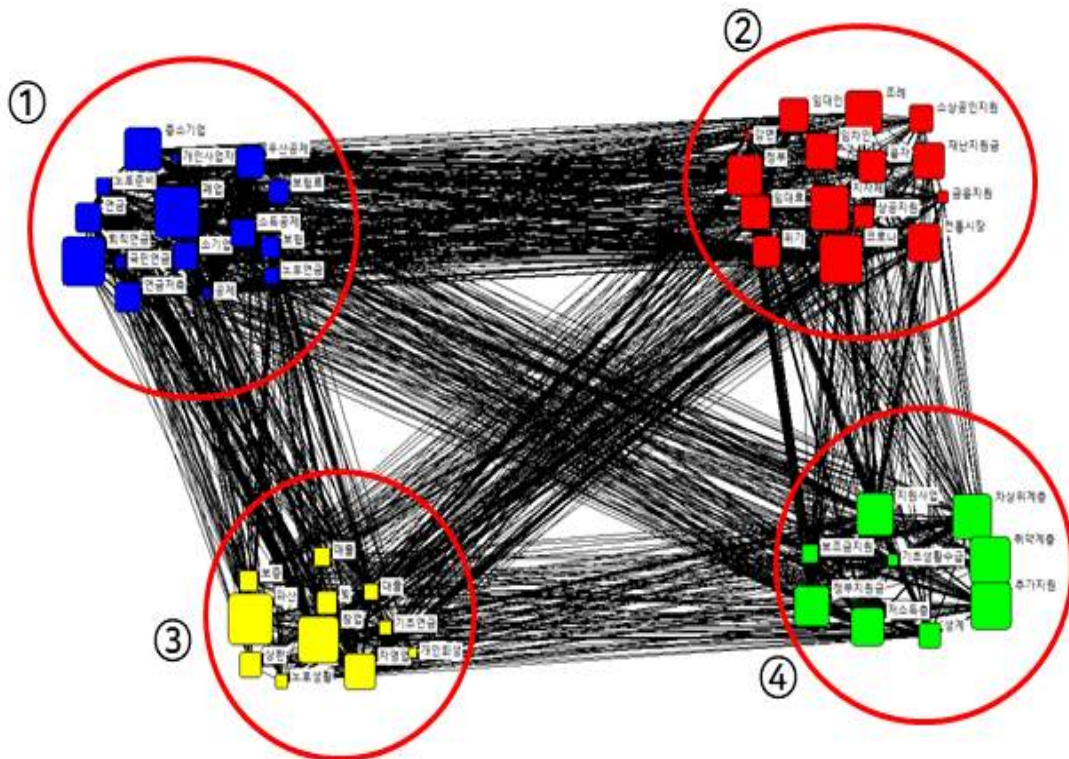
□ 구조적 등위성 분석

- 덴드로그램(Dendrogram)을 통하여 3단계 위계적 집단을 확인
 - 구조적 등위성 분석을 하여 시각화한 결과 <그림 10>과 같고, 키워드 클러스터는 <표 13>과 같음
- 클러스터는 4개가 형성되었고, 각각 다음과 같이 명명함. ① 소상공인의 노후준비에 대한 인식, ② 소상공인의 코로나 지원에 대한 인식, ③ 소상공인의 현 상황에 대한 인식, ④ 소상공인의 정부지원에 대한 인식
 - 그룹 ①은 소상공인들의 노후준비와 관련한 키워드로 구성되어 있고, 그룹 ②는 그룹 ②는 코로나 19 이후 특화된 지원과 관련한 키워드가 포함되어 있음. 가령, 재난지원금, 감면 등이 있음. 그룹 ③은 현 상황에 대한 인식으로 코로나 19 이후 영향을 받았을 것으로 추정되는 매출, 빚, 대출, 등의 키워드

가 포함, 그룹 ④는 정부지원에 대한 인식으로 코로나 19 이후 취약계층으로 전략한 것으로 유추되는 차상위계층, 기초생활수급 등의 키워드가 포함된 것으로 확인

- <그림 >에서 보는 바와 같이 네트워크 연관성은 ①↔②간 가장 높은 것으로 나타나며, 그룹 ①↔②간 일부 키워드의 연관성이 매우 높은 것으로 나타남
 - 가령, ①그룹의 노후연금과 ④그룹의 파산 키워드 간의 연관성이 높으며, ②그룹의 임대료와 ④그룹의 정부지원 키워드 간의 연관성이 높은 것으로 나타남
- 구조적 등위성 분석을 통해 코로나 19 이후 소상공인의 경제적 상황이 악화된 것을 유추할 수 있으며, 코로나 지원이나 정부지원에 의존하는 경향이 높은 것을 확인 할 수 있음

<그림 10> 2기 구조적 등위성 분석 결과 시각화



〈표 13〉 2기 구조적 등위성 분석 주요 키워드 클러스터

	클러스터	핵심단어
1	노후준비에 대한 인식	중소기업, 개인사업자, 우산공제, 노후준비, 폐업, 보험료, 연금, 퇴직연금, 국민연금, 연금저축, 소기업, 공제, 노후연금, 소득공제, 보험
2	코로나 지원에 대한 인식	임대인, 감면, 정부, 임대료, 위기, 코로나, 상공지원, 지자체, 임차인, 용자, 조례, 소상공인지원, 재난지원금, 금융지원, 전통시장
3	현 상황에 대한 인식	보증, 매출, 파산, 빚, 창업, 대출, 기초연금, 자영업, 개인회생, 노후생활, 창업, 상환
4	정부지원에 대한 인식	지원사업, 차상위계층, 취약계층, 추가지원, 기초생활수급, 보조금지원, 정부지원금, 저소득층, 생계

3. 코로나 19 이전과 이후 비교 (1기 vs 2기)

□ 주요 키워드 분석결과, 1기에서는 소상공인 노후준비와 관련된 키워드가 핵심 키워드로 나타났고, 2기에서는 코로나 19의 영향으로 소상공인에 대한 전반적인 지원과 관련된 키워드가 핵심 키워드로 나타남 확인

- 전반적으로 1기에 비해 2기에서 노후준비와 관련된 키워드의 우선순위가 밀리는 것 확인
 - 가령, 노후준비(1기에서는 2위였으나, 2기에서는 5위), 우산공제(1기에서는 9위였으나, 2기에서는 12위) 등의 키워드 순위 하락
- 1기 빈도분석에서 노후준비, 정부, 전통시장, 폐업, 지원사업, 일자리지원 등의 키워드 순위가 높은 것으로 나타났으나, TF-IDF 분석 시 창업, 연금, 우산공제, 노후연금, 퇴직연금 등의 키워드 순위가 높아지는 것 확인
- 2기 빈도분석에서 지원사업, 코로나, 정부, 노후준비, 우산공제, 중소기업 등의 키워드 순위가 높은 것으로 나타났으나, TF-IDF 분석 시 창업, 보조금 지원, 자영업, 보험, 공제, 임대료, 연금 등의 키워드 순위가 높아지는 것 확인

□ 네트워크 분석결과, 1기에 비해 2기에서 소상공인의 경제 상황이 악화된 것 확인. 2기에서 생계, 위기 등이 주요 키워드로 등극

- 이는 코로나 19 이전에는 없거나 순위가 낮았던 키워드가 코로나 19 이후 높은 순위로 등장하는 것에서 유추 가능. 가령, 생계, 폐업, 위기 등의 키워드는 2기에서 중요 키워드로 확인 됨
- 1기 네트워크 분석에서 대출, 노후준비, 창업, 중소기업, 자영업, 정부, 연금, 보험, 소득 등이 연결성 및 영향력이 높은 키워드로 나타났음
- 2기 네트워크 분석에서 코로나, 창업, 정부, 노후준비, 자영업, 중소기업, 대출, 보험, 생계, 폐업, 위기 등이 연결성 및 영향력이 높은 키워드로 나타남

□ 의미연결망 분석결과, 1기에 비해 2기에서 소상공인에 대한 정부지원이 확대 된 것을 확인 할 수 있고, 특히 코로나 지원 등을 통해 소상공인의 정부지원 의존도가 커진 것 확인

- 1기에는 소상공인이 정부지원보다는 개인 대출 등을 주로 이용하여 사업을 설계하는 것으로 나타났고, 노후준비도 현 상황에서 동시 고려하는 것으로 나타남. 2기에서는 소상공인이 코로나 지원이나 정부 지원에 주로 의존하는 것으로 나타났고, 노후준비는 노후준비에 대한 인식은 있으나 현 상황과 동시 고려되지 못하는 것으로 나타남
- 1기 클러스터는 ① 현 상황에 대한 인식, ② 현 사업에 대한 인식, ③ 노후준비에 대한 인식, ④ 정부지원에 대한 인식으로 나타남
 - 소상공인의 현 상황에 대한 인식과 노후준비에 대한 인식의 연관성이 매우 높은 것으로 나타남
 - 소상공인의 경우 정책 지원 보다는 개인 대출 등을 통해 사업 확장, 창업 등을 고려하는 경향이 높은 것을 확인 할 수 있음
 - 또한, 정부지원은 영세소상공인이나 지자체 등을 통해 일자리 지원 사업 등으로 제한적으로 시행되고 있는 것으로 나타남

- 2기 클러스터는 ① 노후준비에 대한 인식, ② 코로나 지원에 대한 인식, ③ 현 상황에 대한 인식, ④ 정부지원에 대한 인식으로 나타남
- 코로나 19 이후 코로나 지원 클러스터가 생겨남으로 소상공인이 개인의 자력 보다는 코로나 지원이나 정부지원에 의존하는 경향이 높음 확인
- 코로나 19 이후 영세소상공인 뿐 아니라 일반 소상공인들까지 지원 범위가 확대된 것 확인. 특히 임대료, 세금 감면 등 정부의 적극적인 지원이 증가한 것을 확인
- 코로나 19 이전에는 취약계층으로 구분할 수 있는 영세소상공인 키워드만 있었다면, 코로나 19 이후에는 차상위계층, 기초생활수급, 저소득층 등도 그룹 ④에 등장하는 것으로 나타나 경제적 상황이 보다 악화되었음을 유추할 수 있음
- 또한 노후준비에 대한 인식은 1기와 2기에서 클러스터로 동일하게 형성되지만 다소 차이가 있는 것으로 확인. 가령, 1기에서는 세액공제, 소득공제 등 개인의 공제 여부가 주요 키워드로 나타나지만, 2기에서는 다소 포괄적인 공제 키워드만 주요하게 나타남

V. 결론: 연구결과 요약 및 시사점

1. 연구결과 요약

□ 주요 키워드 분석결과

- 코로나 19 이전에는 소상공인의 노후준비와 관련된 키워드가 핵심 키워드로 나타났으나, 코로나 19 이후에는 소상공인 지원 관련 키워드가 핵심 키워드로 나타나 코로나 19를 기점으로 소상공인의 노후준비에 대한 인식에 변화가 있었음을 유추할 수 있음

□ 네트워크 분석결과

- 코로나 19 이전보다 코로나 19 이후 소상공인의 경제 상황이 더욱 악화된 것으로 확인됨에 따라 코로나 19를 기점으로 소상공인의 노후준비 경향이 달라졌음을 유추할 수 있음
- 가령, 코로나 19 이후 생계, 위기, 임대인, 감면 등의 키워드가 새롭게 등장하고 실제 네트워크 연결성도 높은 것으로 나타나 소상공인의 현재 경제 상황이 코로나 이전보다 악화되었음을 유추할 수 있음
- 이는 코로나 19로 인해 소상공인 중 일부는 기존의 개인연금이나 노후자금 등을 생계를 위해 이미 모두 사용했거나 사용하고 있다는 신문기사와도 일치하는 것으로 나타남. 이를 통해 코로나 19 이후 소상공인의 노후준비 경향이 다소 달라졌음을 유추할 수 있음

□ 의미연결망 분석결과

- 코로나 19 이전에는 소상공인들이 개인 사업을 하면서 노후준비를 스스로 동시 고려하는 것으로 나타났으나, 코로나 19 이후 노후준비를 포함하여 많은 부분을 소상공인들이 정부지원에 의존하는 것으로 확인
- 1기 클러스터와 2기 클러스터를 비교하면 2기에 코로나 지원과 정부지원이 동시에 등장함으로 소상공인들에 대한 지원 범위가 보다 확대되고 적극적으로 변화했음을 유추할 수 있음
- 또한 1기 클러스터와 2기 클러스터에 모두 등장하는 정부지원에 대한 인식에서 1기에서는 주요 지원 대상에 영세소상공인만 포함되었다면, 2기에서는 차상위계층, 기초생활수급, 저소득층 등 사회취약계층이 다수 포함됨으로써 이들의 경제 상황이 더욱 열악해졌음을 확인할 수 있음. 이를 통해 정부지원을 받는 소상공인이 보다 다양화되고, 양적으로도 증가했음을 유추할 수 있음

□ 연구결과를 바탕으로 종합적으로 분석결과 정리

- 코로나 19 이전과 코로나 19 이후 소상공인의 전반적인 상황이 매우 달라졌고, 이로 인해 노후준비에 대한 인식에도 변화가 있었던 것으로 확인 됨
- 코로나 19 이전 소상공인들은 개인의 여력에 따라 노후준비를 하되, 현업과

노후준비를 동시 고려하는 경향이 높은 것으로 유추할 수 있으나, 코로나 19 이후 어려워진 경제 상황으로 현업과 노후준비를 동시 고려하는 것은 사실상 붕괴된 것으로 유추할 수 있음

- 코로나 19 이후 소상공인의 다수가 노후준비 보다는 현 생계를 위해 정부지원에 의존하는 경향이 높아짐으로써 소상공인의 향후 노후준비도 국가차원에서 고려해야 할 주요 사안으로 인지되고 있음을 유추할 수 있음

2. 시사점

□ 소상공인의 노후준비에 대한 정부의 인식 전환 필요

- 기존 소상공인의 노후준비는 개인의 문제로 여기는 경향이 높았으나, 코로나 19를 겪으면서 소상공인의 폐업과 은퇴는 단순히 개인의 문제가 아니라 심각한 사회문제가 될 수 있음을 인지함. 따라서 소상공인 노후준비에 대한 정부의 보다 적극적이고 유연한 인식 전환 필요
- 코로나 19 이후 소상공인 간 노후준비에 대한 격차가 커졌을 것으로 유추되므로 사회적 균형을 맞추기 위한 정부의 다양한 지원책 마련이 필요
- 다만, 장기적인 관점에서 국민 간 형평성의 문제 및 복지재정의 부담 등을 야기할 수 있으므로 지원범위를 조절 할 필요가 있음

□ 소상공인의 노후준비에 대한 대국민적 공익 홍보 강화 필요

- 코로나 19 이후 경제 상황이 악화되면서 소상공인들이 스스로 노후준비를 포기하는 경향이 높아지고, 코로나 지원이나 정부보조금의 확대에 정부지원에 의존하는 경향이 높아진 것으로 확인
- 이는 지속가능한 소상공인 지원을 방해할 수 있으므로 장기적인 관점에서 소상공인 스스로 노후준비의 필요성을 인지하고, 준비할 수 있도록 홍보하는 것이 필요
- 또한 국민 누구나 소상공인이 될 수 있으므로 대상을 특정하지 않고, 공익 홍보를 강화함으로써 소상공인의 노후준비에 대한 전국민의 관심을 모으는 것이 필요하며 이를 통해 소상공인의 노후준비에 대한 사회적 인식 변화를

도모할 수 있을 것임

□ 소상공인의 노후준비를 도울 수 있는 다양한 금융지원책의 보완 필요

- 현재 소상공인 지원을 위한 다양한 지원기관이 있고, 금융 상품이 있음
 - 가령, 한국주택금융공사에서는 영세자영업자 전세자금보증(특례), 주택담보대출 원금상환 유예제도, 유동화미수채권 분할상환금 상환 유예제도, 주택보증구상채권 분할상환금 상환유예제도 등을 시행함으로써 소상공인을 지원하고 있음
- 그러나 근본적으로 소상공인의 현 상황을 도와주는 것 외에 노후준비를 위한 직접적인 상품은 노란우산공제 상품이 거의 유일함. 따라서 관련 상품의 다양화를 통해 소상공인이 노후준비를 할 수 있는 다양한 선택지가 마련될 필요가 있음

□ 장기적인 관점에서 소상공인의 지속가능한 지원책으로 주택연금 연계 필요

- 주택연금 제도 개선을 통해 소상공인 가입자 확대 및 지원을 고려할 필요가 있음
 - 현재 소상공인이 주거지와 사업장을 겸해서 사용하고 있는 복합용도주택(근린생활시설)이 주택연금 가입요건에 해당되기 위해서는 주택으로 사용하는 면적이 건물 전체 면적의 50% 이상이어야 하고, 주택연금이 가능하더라도 연금지급액이 주택 부분에서만 해당되어 연금액 산정 시 주택 부분만 반영하고 상업용 등 다른 용도 부분은 제외됨. 또한 연금액은 주택부분에서만 나오는데 저당권 설정은 주택이 아닌 부분까지 포함한 전체가 해당 됨
 - 따라서 제도 개선을 통해 소상공인이 필요하다면 조건을 보다 완화 적용할 수 있는 방안을 신중히 고려할 필요가 있음
- 현재 공사에서 추진 중인 소상공인 주택연금 연계 상품의 활성화 필요
 - 현재 공사는 「서민 등 취약계층 맞춤형 금융지원 확대 방안」의 후속조치로 폐업 희망 소상공인이 자가 주택에서 계속 거주하면서 개인사업자 대출 등을 상환할 수 있는 주택연금 상품 신설 중임
 - 주요 내용은 주택연금 신청시점 기준으로 「소상공인기본법」상 소상공인 요건

을 충족하는 개인사업자 대상으로 기존 대출 상환이 가능한 상품을 신설하고, 개인사업자 대출 상환용 주택연금 상품에 가입 후 6개월 내 폐업신고 미이행 시 대출 회수함

- 이와 같이 새로운 상품의 출시를 통해 소상공인들이 현 상황의 어려움을 극복함과 동시에 노후준비를 할 수 있도록 할 수 있는 적극적인 노력 필요

참 고 문 헌

- 강용희(2021), Textom과 함께하는 빅데이터 아카데미 강의자료
- 금융위원회(2014), 한국주택금융공사법 시행령 일부개정안 조문별 제·개정이유서
- 김수환·김형규(2017), 소상공인의 사회안전망에 관한 연구, 전문경영인연구, 20(1), pp.69-87
- 김윤영(2020), 불안정 노동자의 노후소득보장: 연금유형에 따른 불안정 노동자의 연금소득 시뮬레이션 분석, 한국사회정책, 27(2), pp.163-195
- 문지영·김학선·이종호(2020), 의미연결망 분석을 활용한 음식관광 활성화 방안에 관한 연구, 산업혁신연구, 36(2), pp.135-153
- 문현경·김혜진(2023), 연금 개혁의 주요 행위자로서 사용자에 대한 기초연구, 정부학연구, 29(3), pp.43-83
- 성미영·노미나·서병태, 텍스트마이닝 기법을 활용한 아동 재난안전 정보에 대한 사회적 인식 분석, 아동과권리, 25(2), pp.167-185
- 이광일·남정민(2024), 텍스트 마이닝 기법을 활용한 코로나-19 팬데믹 발생 이전과 이후 스타트업 성장요인에 관한 인식 변화 연구, 한국진로창업경영학회지, 8(1), pp.19-38
- 이용규(2022), SNS빅데이터를 활용한 코로나-19 팬데믹 발생 전, 후의 여행 마케팅 타겟의 트렌드 변화 연구, 마케팅논집, 30(2), pp.30-49
- 이재무·박인식(2021), 텍스트마이닝을 활용한 사회적기업에 대한 사회적 인식 분석, 사회적가치와 기업연구, 14(2), pp.111-135
- 이지백(2024), 텍스트마이닝을 활용한 부산 부동산 시장 인식 변화 연구, 인문사회과학연구, 32(4), pp.685-713
- 전병욱(2016), 노후보장 저축제도의 과세제도 개선방안, 예산정책연구, 5(1), pp.50-80
- 전병욱(2017), 소기업·소상공인 공제제도의 문제점과 개선방안, 경영법률, 27(2), pp.43-775
- 정수원(2013), 일본 소규모기업의 공제제도 현황과 과제, 일본근대화연구, 40, pp.405-430

- 중소기업중앙회(2019), 소상공인 경영애로 실태 및 정책과제 조사 결과보고서, pp.1-25
- 중소기업중앙회(2021), 코로나 19로 인한 소상공인 일(Work)과 삶(Life)의 변화 조사 보고서, pp.1-169
- 최윤정·권상희(2014), 빅데이터 관련 신문기사의 의미연결망 분석, 사이버커뮤니케이션 학보, 31(1), pp.241-286
- 홍희정(2022), 빅데이터 분석을 통한 초장기 정책모기지에 대한 사회적 인식 연구, 한국웰니스학회지, 17(4), pp.29-36
- Bhargavi, P., Jyothi, B., S. and Sekar, K. (2008), Knowledge extraction using rule based decision tree approach, IJCSNS International Journal of Computer Science and Network Security, 8(7), pp.296-301
- Coombes, P. J., & Barry, M. E. (2014). "A systems framework of big data driving policy making-Melbourne's water future." In OzWater14 Conference. Brisbane: Australian Water Association.
- Fan, J., Han, F., & Liu, H. (2014). "Challenges of big data analysis". National science review. 1(2). pp. 293-314
- Freeman, L. C. (1979). Centrality in Social Networks Conceptual Clarification. Social Networks, 1(3), pp.215-239
- Racherla, P., & Hu, C. (2009). A framework for knowledge based crisis management in the hospitality and tourism industry, Crnell Hospitality Quarterly, 50(4), pp.561-577
- Sagiroglu, S., & Sinanc, D. (2013). "Big data: A review". In 2013 International Conference on Collaboration Technologies and Systems (CTS). pp.42-47. IEEE.
- Sehgal, D., & Agarwal, A. K. (2018). Real-time sentiment analysis of big data applications using twitter data with Hadoop framework. Advances in Intelligent Systems and Computing, 2(1), pp.765-772
- Ward, J. S., & Barker, A. (2013). Undefined by data: a survey of big data definitions. arXiv preprint arXiv.
- 국제신문(2024.5.19.) "임대료. 빛에 허덕여... '환갑의 사장님'들 노후자금 갠다, kookje.co.kr/news2011/asp/newsbody.asp?key=20240520.22012005235

한국주택금융공사 홈페이지, <https://hf.go.kr>

한국학술지인용색인 홈페이지, www.kci.go.kr