



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

석 사 학 위 논 문

플랫폼 데이터를 활용한
비프랜차이즈 카페 매출 영향 요인 연구

숙명여자대학교 대학원

경영학과 비즈니스애널리틱스 전공

송 혜 민

석 사 학 위 논 문

플랫폼 데이터를 활용한
비프랜차이즈 카페 매출 영향 요인 연구

숙명여자대학교 대학원

경영학과 비즈니스애널리틱스 전공

송 혜 민

석사학위논문

플랫폼 데이터를 활용한
비프랜차이즈 카페 매출 영향 요인 연구

지도교수 전 미 나

이 논문을 경영학 석사 학위 논문으로 제출함

2024년 6월

숙명여자대학교 대학원

경영학과 비즈니스애널리틱스 전공

송혜민

송혜민의 경영학 석사 학위 청구 논문을 인준함

플랫폼 데이터를 활용한
비프랜차이즈 카페 매출 영향 요인 연구

2024년 6월

심사위원장 (인)

위 원 (인)

위 원 (인)

숙명여자대학교 대학원

목 차

목 차	i
표 목 차	iv
그림 목차	vi
국문 초록	viii
I. 서론	1
1. 연구 배경 및 목적	1
2. 상권 별 현황 조사	5
II. 문헌 연구 및 연구 개요	6
1. 외식업 상권 연구	6
2. 카페 선택 속성 연구	7
3. 외식업 리뷰 연구	8
4. 연구 개요	9
III. 연구 1: 매출에 영향을 주는 입지 요인에 대한 분석	10
1. 연구 방법	10
1.1 연구 대상	10

1.2 자료 수집	11
1.3 변수	12
1.3.1 종속변수	12
1.3.2 독립변수	12
2. 연구 결과	16
2.1 오피스 상권 기초통계와 상관관계	16
2.2 오피스 상권 회귀 분석	18
2.3 핫플레이스 상권 기초통계와 상관관계	20
2.4 핫플레이스 상권 회귀 분석	22
 IV. 연구 2: 매출 상위 카페 리뷰 분석	24
1. 연구 방법	24
1.1 연구 대상	24
1.2 자료 수집	24
1.3 분석 절차	26
2. 연구 결과	28
2.1 오피스 상권 TF 및 TF-IDF 분석 결과	28
2.2 핫플레이스 상권 TF 및 TF-IDF 분석 결과	32
2.3 워드 클라우드	36
2.4 토픽모델링 결과	38

V. 추가 분석 1: 핫플레이스 상권 중가 프랜차이즈 리뷰 분석	46
1. 연구 방법	46
1.1 연구 대상	46
1.2 자료 수집	46
2. 연구 결과	47
2.1 TF 및 TF-IDF 분석 결과 & 워드 클라우드	47
VI. 추가 분석 2: 작성 회차에 따른 오피스 상권 리뷰 비교	51
1. 연구 방법	51
1.1 연구 대상	51
1.2 자료 수집	51
2. 연구 결과	52
2.1 리뷰 작성 회차에 따른 TF 순위 비교	52
VII. 결론	55
1. 연구 결과 요약	55
2. 학문적 시사점	57
3. 실무적 시사점	58
4. 한계점 및 향후 연구 방향	59
VIII. 참고 문헌	60
ABSTRACT	63

표 목 차

<표 1> 오피스 상권 기초통계량	17
<표 2> 오피스 상권 회귀 분석 결과	19
<표 3> 핫플레이스 상권 기초통계량	21
<표 4> 핫플레이스 상권 회귀 분석 결과	22
<표 5> 상권 별 수집된 리뷰 수	25
<표 6> 오피스 상권 TF 분석 결과	30
<표 7> 오피스 상권 TF-IDF 분석 결과	31

<표 8> 핫플레이스 상권 TF 분석 결과	34
<표 9> 핫플레이스 상권 TF-IDF 분석 결과	35
<표 10> 오피스 상권 토픽 모델링 결과	43
<표 11> 핫플레이스 상권 토픽 모델링 결과	45
<표 12> 핫플레이스 상권 증가 프랜차이즈 TF 분석 결과	48
<표 13> 핫플레이스 상권 증가 프랜차이즈 TF-IDF 분석 결과	49
<표 14> 작성 회차 별 수집된 리뷰 수	51
<표 15> 리뷰 작성 회차에 따른 오피스 상권 TF 분석 결과	53

그림 목 차

[그림 1] 주요 지도 앱 2023년 월 평균 사용자 수	3
[그림 2] QGIS 활용 분석을 통한 독립 변수 획득 과정	15
[그림 3] 회귀 분석 연구 모형	16
[그림 4] 외식산업경기동향지수 추이	25
[그림 5] 오피스 상권 워드 클라우드	36
[그림 6] 핫플레이스 상권 워드 클라우드	37
[그림 7] 오피스 상권 perplexity & coherence	40
[그림 8] 핫플레이스 상권 perplexity & coherence	40

[그림 9] 핫플레이스 상권 증가 프랜차이즈 워드 클라우드	50
--	----

국 문 초 록

카페 시장의 포화에 따라 카페 폐업 수가 급증하고 있다. 이러한 상황에서 비프랜차이즈 카페들은 저가 프랜차이즈 카페의 가격 경쟁력과 고가 프랜차이즈 카페의 자본에 의해 이중고를 겪고 있다. 이에 본 연구는 핫플레이스 및 오피스 상권의 비프랜차이즈 카페를 대상으로 두 가지 탐색적 연구를 통해 프랜차이즈 카페 출점 제한 제도 도입 및 비프랜차이즈 카페 포지셔닝 전략 수립에 대한 시사점을 제공하고자 한다. 이를 위해 빅데이터 상권 분석 플랫폼 오픈업에서 추정 매출 데이터를, 네이버 지도에서 리뷰 데이터를 수집하였다. 연구 1에서는 주변 카페의 가격대와 수가 추정 매출에 미치는 영향을 다중회귀분석으로 살펴보았다. 그 결과, 두 상권 모두에서 주변에 고가 프랜차이즈 카페가 많을수록 제곱 미터 당 평균 추정 매출이 증가하였으며, 핫플레이스 상권에서는 주변의 증가 프랜차이즈 카페 수가 증가할수록 제곱 미터 당 평균 추정 매출이 감소하는 것으로 나타났다. 연구 2에서는 상권별 제곱 미터 당 평균 추정 매출 상위 20개 카페 리뷰에 대한 LDA 토픽모델링을 통해 두 상권의 특성 차이를 확인하였다. 그 결과, 오피스 상권 및 핫플레이스 상권의 비프랜차이즈 카페는 모두 커피 구매, 디저트 구매, 카페 탐방을 목적으로 하는 세 개의 군집으로 구분되었으며, 구체적인 내용에는 상권 특성에 따른 차이가 존재하는 것으로 나타났다.

주제어 : 비프랜차이즈 카페, 온라인 리뷰, 토픽모델링, 텍스트마이닝

I . 서론

1. 연구 배경 및 연구 목적

2023 년 유로모니터의 조사 결과에 따르면 한국인 1 인당 연간 커피 소비량은 405 잔이다. 이는 전 세계 152 잔의 2 배 이상이며, 미국보다 80 잔 많은 수치이다. 이처럼 현재 대한민국은 커피 공화국이라고 칭할 수 있을 만큼 높은 커피 소비량을 보이고 있다.

커피의 뜨거운 인기에 힘입어 프랜차이즈 카페 수가 크게 증가하고 있다. 공정거래위원회에 따르면 2022 년 커피 전문점 프랜차이즈 브랜드 수는 852 개로, 2 년 전인 2020 년에 비해 두 배 넘게 증가하였다.

카페 수가 늘어남에 따라 폐업하는 카페 수 역시 크게 증가하는 문제가 발생하고 있다. 행정안전부 인허가 데이터에 따르면 폐업한 카페의 수는 2023 년에만 1 만 1450 개로, 2020 년 9111 개에 비해 크게 증가하였다. 전국 커피전문점 수가 2022 년 말 기준 10 만 729 개라는 점을 고려할 때(통계청, 2022 서비스업조사) 2023 년 한 해에만 약 10%의 카페가 폐업했음을 시사한다.

카페 폐업이 증가하는 문제에 대한 첫 번째 이유로 카페 소비의 양극화 현상이 있다. 스타벅스, 투썸플레이스 등 가격대가 높은 고가 프랜차이즈들은 높은 자본을 바탕으로 소비자들의 가격 대비 심리적 만족도, 이른바 ‘가심비’를 만족시킨다. 최근 그 수가 압도적으로 증가하고 있는 메가 MGC 커피, 컴포즈커피, 뽕다방 등 저가 프랜차이즈들은 낮은 지갑을 지닌 소비자들의 가격 대비 성능에 대한 만족도, 이른바 ‘가성비’를 만족시킨다. 이러한 상황에서 비프랜차이즈 카페들은 프랜차이즈 카페에 밀려 이중고를 겪고 있다. 원두 구매 과정에서 규모의 경제를 따라잡기 어려워 저가 프랜차이즈의 가격 경쟁력에 밀리고, 고가 프랜차이즈 브랜드의 자본력과 고급 브랜드 이미지에 밀려 경쟁력을 잃고 있다.

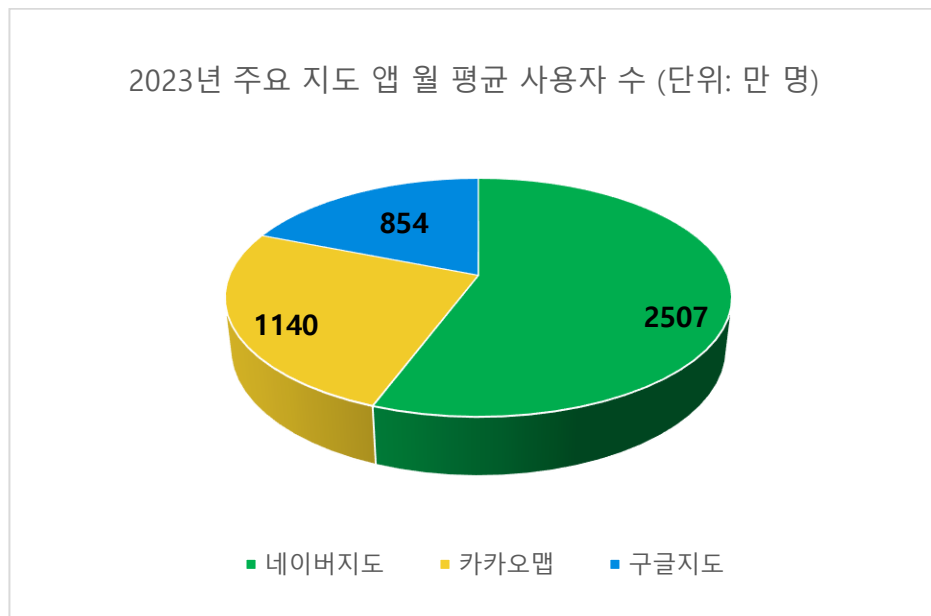
카페 폐업이 증가하는 두 번째 이유로 출점 규제의 부재를 들 수 있다. 담배 판매권을 기준으로 출점이 제한되는 편의점 프랜차이즈, 배달 권역을 기준으로 동일 프랜차이즈 업체의 출점이 제한되는 치킨 프랜차이즈와 달리 카페는 배달 중심이 아니기 때문에 권역 설정이 어렵다. 출점 규제 없이 무분별한 개업이 이루어지고 있다. 이를 해결하기 위한 제도적 노력이 없었던 것은 아니다. 2012년 공정거래위원회는 프랜차이즈 카페의 무분별한 출점을 막기 위해 ‘모범 거래 기준’을 도입하여 동일 브랜드 가맹점 출점 시 반경 500M 내 신규 출점을 금지하였으나 2014년 폐지하였다. 스타벅스처럼 가맹점이 아닌 직영점 형태로 운영되는 카페들은 규제를 피해간다는 지적이 나왔기 때문이다(머니투데이, 2023).

비프랜차이즈 카페들은 경쟁력 강화를 위해 다방면으로 노력을 기울이고 있다. 특히 업체 관리 플랫폼을 활용하여 온라인에서 잠재 고객을 유인하고 있다. 배달을 하는 카페의 경우, 배달 플랫폼인 배달의 민족, 요기요, 쿠팡이츠에 카페에 대한 정보를 제공한다. 소셜미디어를 활용하여 카페를

마케팅하기도 한다. 많은 비프랜차이즈 카페들이 유튜브 채널을 운영하거나 인스타그램, 페이스북 계정을 운영하며 신메뉴 소개, 운영 안내 사항 등을 업로드함으로써 고객과 소통하고 있다. 지도 플랫폼인 네이버 지도, 구글 지도, 카카오 맵에도 카페 정보를 세팅해 매장을 검색하는 고객을 유인하고 있다.

이처럼 다양한 업체 관리 플랫폼 중에서도 네이버 지도는 대표적인 비프랜차이즈 업체 관리 플랫폼이다. 네이버 지도는 2023 년을 기준으로 지도 앱 중 점유율 1위를 차지하고 있다(와이즈 앱, 2023). 또한 무료로 예약 시스템을 활용할 수 있으며, 배달의 민족과의 연동을 통해 배달 서비스를 제공하는 등 종합 플랫폼의 성격을 지닌다.

[그림 1] 주요 지도 앱 2023년 월 평균 사용자 수 (데이터 출처: 와이즈앱)



사업주는 네이버 스마트 플레이스를 활용해 네이버 지도 검색에 업체를 등록할 수 있고, 검색되는 업체 정보를 직접 관리할 수 있다. 업체 정보의 예시로는 업체 이미지, 새소식, 영업 시간, 신메뉴, 리뷰 등이 있다. 방문자 리뷰의 경우, 영수증이나 네이버 예약 등으로 장소를 방문한 사실을 인증한 사용자만 작성이 가능하다. 이에 따라 실이용자의 진성 후기가 제공된다는 장점이 있다.

한편, 상권 분석 플랫폼 오픈업은 여러 데이터를 학습한 AI 를 통한 빅데이터 분석으로 개별 매장 별 추정 매출 정보, 유동 인구 특성 등의 상권 정보를 제공하고 있다. 해당 플랫폼을 통해 사업주들은 입지 선정 시 오픈업의 데이터를 활용해 의사 결정에 도움을 얻을 수 있다.

이에 따라 본 연구는 대표적인 업체 관리 플랫폼인 네이버 지도에 나타난 리뷰 정보와 상권 분석 플랫폼 오픈업에 나타난 추정 매출 정보를 바탕으로 오피스 상권과 핫플레이스 상권에 위치한 비프랜차이즈 카페들의 특성을 연구하고자 한다. 이를 통해 두 상권에서 비프랜차이즈 카페의 경쟁력을 높일 수 있는 포지셔닝 전략을 탐구하고자 한다.

2. 상권 별 현황 조사

2.1. 역삼 1 동 (행정동)

강남구의 중앙 서쪽에 위치한 역삼 1 동은 강남대로와 2 호선 강남역이 교차하여 국내에서 유동 인구가 가장 많은 상권(한국데이터산업진흥원, 2020)이다. 1966 년 서울 도시기본계획을 통해 본격적으로 개발되기 시작했다. 계획을 통해 개발된 도시 특성상 봉은사로, 선릉로, 논현로, 언주로, 역삼로, 테헤란로 등이 바둑판식으로 배열되어 교통이 매우 편리하다.

2.2. 서교동 (행정동)

마포구의 중앙부에 위치한 서교동은 국내에서 유동인구가 세 번째로 많은 상권으로(한국데이터산업진흥원, 2020), 홍대앞 상권에 해당한다. 과거 상수동, 서교동 일대를 가리키던 홍대앞 상권은 현재 합정동, 연남동, 망원동, 성산동 일대까지 확장된 영역으로 간주되고 있다.

홍대앞 상권은 1970 년대부터 홍익대학교 미술대학생들의 작업실 문화를 기반으로 독창적이면서도 다채로운 장소성을 가지게 되었다. 1990 년대 초 고급 카페 문화, 1990 년대 말 클럽 문화를 형성하며 다수의 젊은 층이 방문하는 거리가 되었다. 이후 대규모 자본의 유입과 지하철 개통으로 유동인구가 상승하며 상업화되었다(서울특별시, 2017).

II. 문헌 연구 및 연구 개요

1. 외식업 상권 연구

프랜차이즈 외식업체의 증가와 함께, 프랜차이즈와 비프랜차이즈 외식업체 비교에 대한 여러 선행 연구가 이루어졌다. 최누리와 김인신(2022)은 부산 카페의 생존함수를 추정하여 부산시 카페 상권의 상업젠트리피케이션 현황을 진단하였다. 분석 결과 전포와 영도 카페 상권에 위치한 카페의 경우 프랜차이즈 카페의 생존기간이 독립카페보다 높게 유지되었다. 이에 따라 도시계획차원에서 상권 특성을 반영한 적극적인 보호 및 지원 정책이 필요성을 입증하였다.

본 연구의 대상인 핫플레이스 상권에 대한 연구도 수행되어 왔다. 민철기·강창덕(2021)는 대표적인 핫플레이스 상권인 홍대지역 음식점의 상업시설 생존율을 비교하였다. 그 결과 홍대앞상권에서는 입지 특성이나 인구 특성, 경제 특성보다는 점포의 특성이 생존에 미치는 영향이 크게 나타났다. 특히 프랜차이즈 음식점의 폐업 위험이 비프랜차이즈 음식점보다 낮게 나타나며 프랜차이즈 변수가 폐업 위험에 매우 중요한 상대적 영향을 미쳤다. 또한 김성아 외(2023)에 따르면 2022 년에 핫플레이스 형성에 유의미하게 작용한 변수는 프랜차이즈 수, 아파트 주거 연면적, 토지이용복합도로 나타났다. 즉 선행연구를 통해 프랜차이즈 수가 핫플레이스 형성에 유의미한 영향을 미침을 확인할 수 있었다.

2. 카페 선택 속성 연구

신재욱(2014)는 프랜차이즈 커피 브랜드의 차별화 전략을 연구하였다. 분야별 STP 분석 결과 맛보다 가격을 중시하는 ‘가격 중시형’, 가격과 맛을 동시에 추구하는 ‘조화형’, 개성표현의 한 형태로 커피를 생각하는 ‘개성 추구형’, 맛을 가장 우선으로 생각하는 ‘매니아형’, 편안한 공간을 가장 중요시하는 ‘공간중시형’으로 분류되었다.

배준호, 이준재, 박정숙(2014)은 카페의 선택속성을 도출하기 위해 강릉 커피거리 커피전문점의 21 개의 속성을 토대로 요인분석을 실시하였다. 그 결과 유형성, 쾌적성, 신뢰성, 반응성, 보증성의 5 가지 요인이 도출되었다.

최성임, 임은순, 문혜선(2012)은 커피 선택에 영향을 미치는 요인을 조사하였다. 그 결과 직업, 나이, 포장여부에 따라 커피 메뉴 선택이 달라졌다. 또한 ‘기분을 좋게해서’, ‘멋져보여서’, ‘체중조절을 위해’, ‘습관적으로’ 커피 메뉴를 선택한다는 결과가 나타났다.

임병학과 엄성원(2022)은 카페 리뷰 감성 분석에 SERVQUAL모형을 적용하였다. 구글 맵에 나타난 카페 리뷰에 감성 분석을 통해 문장 단위로 감성 점수를 부여하고, 문장 별 키워드를 SERVQUAL모형의 각 차원으로 분류하여 차원 별 점수를 합산하였다. 이를 통해 각 차원에 대한 카페의 서비스 품질 점수를 측정하였다. 그 결과, 카페 서비스 품질 점수는 확신성과 유형성 순으로 높게 나타났으며, 반응성이 가장 낮게 나타났다. 유형성의 예시로는 ‘인테리어’, ‘위치’, ‘깔끔’ 등이 해당되었다.

3. 외식업 리뷰 연구

외식업 리뷰가 매출에 미치는 영향에 관한 연구로 김용원 외(2021)는 연구를 통해 서울시 법정동 내 음식점의 평균 매출과 블로그 게시글 수는 서로 정(+)의 관계에 있음을 밝혔다. 이는 지역별 데이터를 활용한 연구로, 연구자는 연구의 한계점으로 블로그 정보의 가치를 마케팅 측면에서 정량적으로 평가하기 위해 지역별 데이터가 아닌 개별 음식점 데이터를 활용하여야 한다고 언급하였다.

네이버 지도와 연동되는 네이버 플레이스가 대표적인 외식 업체 관리 플랫폼으로 자리매김하면서 네이버 플레이스에서 수집한 데이터를 활용한 여러 연구가 진행되었다. 윤동화(2022)는 네이버 플레이스에 나타난 2354개의 제주 음식점 키워드 리뷰 데이터를 K-Means 군집분석하였다. 그 결과, 동지역이 읍면지역보다 공간적인 면과 가성비적 측면에서 더 높은 만족도를 보였다.

정동영 외(2022)에 따르면 네이버 플레이스에서 기존 고객의 리뷰가 신규고객의 리뷰보다 긍정적으로 나타났다. 또한 신규 고객과 기존 고객의 리뷰 감정은 업체 페이지 유입에 긍정적인 영향을 보였다. 주목할 만한 지점은 업체의 성과와 관련된 전화걸기, 길찾기, 공유와 같은 변수들은 기존 고객 리뷰에서만 긍정적인 영향을 주었다는 것이다. 즉 네이버 플레이스에서 신규 고객의 리뷰보다 기존 고객의 리뷰가 소비자의 구매 전환과 업체의 매출 증대에 효과적이라는 사실을 도출하였다.

4. 연구 개요

본 연구의 연구 문제는 다음과 같다.

[연구 문제 1] 주변 프랜차이즈 카페의 가격대와 수는 비프랜차이즈 카페의 평균 추정 매출에 영향을 미치는가?

[연구 문제 2] 상권의 종류에 따라 상위 매출 카페의 리뷰 특성은 어떻게 군집화되는가?

본 연구의 목적은 비프랜차이즈 카페의 경쟁력 강화를 위한 시사점을 도출하기 위해 오피스 상권과 핫플레이스 상권에 위치한 비프랜차이즈 카페의 특징을 도출하고 비교하는 데 있다.

구체적으로 연구 1에서 경쟁 관계에 있는 주변 프랜차이즈 카페들과의 비교를 통해 각 상권 비프랜차이즈 카페의 입지에 따른 매출에의 영향을 도출하고, 연구 2에서는 각 상권 별 상위 매출 카페의 리뷰 특성을 탐색함으로써 각 상권의 비프랜차이즈 카페의 포지셔닝 전략에 적용 가능한 시사점을 도출하고자 한다.

Ⅲ. 매출에 영향을 주는 입지 요인에 대한 분석

1. 연구 방법

1.1. 연구 대상

분석 대상 행정동으로는 핫플레이스 상권과 오피스 상권을 비교하기 위해 오피스 상권으로 역삼 1 동을, 핫플레이스 상권으로 서교동을 선정하였다. 역삼 1 동은 대형 기업들의 업무 시설이 높게 밀집해 있는 국내 대표 오피스 상권인 역삼역 상권을 포함한다. 서교동은 홍대 상권을 아우르는 행정동으로, 홍대는 서울의 핫플레이스로 다수 연구에서 언급되고 있다(정의석 외, 2020). 이에 더해 두 행정동은 서울시 상권분석서비스의 ‘뜨는 상권’에서 2023 년 4 분기를 기준으로 커피/음료 분야 점포 수 상위 2 개 행정동이자, 매출 상위 5 개 행정동에 속한다. 이에 따라 타 행정동에 비해 충분한 비프랜차이즈 카페 수를 보유하고 있을 것으로 유추된다.

연구 대상 카페 후보 목록은 다음과 같이 생성하였다. 여러 연구에서 활용되고 있는 행정안전부의 지방행정인허가데이터 중 휴게음식점 데이터를 활용하였고, 네이버지도에서 ‘역삼 1 동 카페’, ‘서교동 카페’로 검색되는 카페들 중 중복되지 않는 카페들을 추가로 수집하였다.

연구 대상 카페 선정에는 세 가지 기준을 활용하였다. 첫 번째로 QGIS 를 활용하여 행정동 경계를 기준으로 300M 내에 있는 카페들만을 대상으로 선정하였다. 이는 기준 데이터 수를 늘리기 위한 것으로, 고객이 카페 접근에

도보를 선호하여 도보 5분거리에 해당하는 매장 반경 300m의 특성이 매출에 유의미하다는 선행 연구(신우진 외, 2011)에 따른 것이다. QGIS는 안예린 외(2023)가 머신러닝을 활용한 커피 프랜차이즈 실패 요인 연구에서 이용한 지리 시각화 툴로, 다수의 상권 분석 연구에서 활용되고 있다.

두 번째로 공정거래위원회 가맹사업거래 사이트에서 정보공개서가 검색이 되지 않는 비프랜차이즈 카페만을 대상으로 하였다. 스타벅스 역시 모두 가맹이 아닌 직영점으로 운영되나 소상공인을 대상으로 시사점을 내기 위한 연구의 특성상 연구 대상 비프랜차이즈 카페 리스트에서 제외하였다. 마지막으로 오픈업에 추정 매출 데이터가 공개된 카페만을 대상으로 하였다. 이에 따라 오피스 상권인 역삼 1동의 비프랜차이즈 카페 114 곳과 핫플레이스 상권인 서교동의 비프랜차이즈 카페 105 곳을 연구 대상으로 선정하였다.

1.2 자료 수집

본 연구를 위해 수집한 데이터는 크게 세 종류로 나뉜다. 첫 번째는 오픈업에 나타난 개별 매장의 추정 매출 데이터이다. 두 번째로 연구 대상 카페 선정에 활용했던 지방행정인허가데이터의 ‘소재지 면적’ 칼럼을 활용하였다.

연구는 비프랜차이즈 카페 추정매출에 영향을 미치는 다양한 요인들을 살펴보기 위해 빅데이터 상권 분석 플랫폼 오픈업의 데이터를 활용하였다. 기존의 카페 매출에 대한 연구의 대다수는 공공데이터에 나타난 블록 단위 별 추정 매출을 기준으로 행해졌으나, 이들 연구가 한계점에서 언급하였듯 개별 매장에 대한 연구가 이루어질 필요가 있을 것으로 판단된다(김용원외, 2021).

이에 본 연구는 국가 부처, 통신사, 카드사 등에서 가져온 다양한 빅데이터를 토대로 인공지능이 계산한 오픈업의 개별 매장 별 추정 매출을 조사하였다.

1.3 변수

1.3.1 종속변수

본 연구에서는 오픈업에 나타난 2024 년 4 월 개별 매장의 추정 매출 데이터를 종속변수 생성에 활용하였다. 오픈업에 매장 명을 검색하면 최저 추정 매출과 최대 추정 매출이 나타난다. 이 둘을 모두 조사하여 평균치를 내어 ‘평균 추정 매출’ 을 구하고, 이를 지방행정인허가데이터에 존재하는 ‘소재지 면적’ 으로 나누어 ‘제곱 미터 당 평균 추정 매출’ 이라는 종속변수를 생성하였다.

오픈업 매출은 실제 수치가 아닌, 국가 부처·통신사·카드사 데이터 등 다양한 빅데이터를 학습하여 AI 가 계산한 추정값이다. 이에 따라 추정 매출이라 명명하였다.

1.3.2 독립변수

본 연구의 독립변수로 활용된 ‘300M 내 고가 프랜차이즈 카페 수’ , ‘300M 내 중가 프랜차이즈 카페 수’ , ‘300M 내 저가 프랜차이즈 카페 수’ , ‘300M 내 고가 비프랜차이즈 카페 수’ , ‘300M 내 중저가 비프랜차이즈

카페 수' 는 QGIS 를 활용하여 측정하였다.

독립변수를 획득하기 위해 QGIS 프로그램을 활용하였다. 독립변수 획득 과정은 다음과 같다. 우선 연구 대상 데이터가 행정동 기준 300M 내에 위치한 비프랜차이즈 카페임을 고려해, 행정동 기준 300M 밖으로부터 또다시 주변 300M 내에 있는 카페 리스트를 획득하고자 하였다. 이에 따라 행정안전부 인허가데이터 및 네이버 검색으로 획득한 카페 리스트에서 각 행정동 주변 600M 내에 위치한 카페 데이터를 추려내었다.

이때에는 연구 대상에서 제외했던 스타벅스를 고가 프랜차이즈 카페 리스트에 포함하였다. 앞서 언급하였듯, 직영점으로 운영되는 스타벅스가 모범 거래 기준 대상에서 제외된다는 반발로 인해 모범 거래 기준 제도가 폐지되었다. 이는 스타벅스가 직영점으로 운영됨에도 불구하고 프랜차이즈 점포 출점 제한의 대상이 되어야 한다는 사회적 인식이 지배적임을 의미한다. 이에 따라 소상공인을 위한 시사점을 내리는 연구의 특성상 스타벅스를 고가 프랜차이즈 리스트에 포함하였다.

다음으로 데이터 시트를 QGIS 로 불러와 각 행정동 shp 파일을 기준으로 600M 버퍼를 생성한 후, '위치로 선택' 의 '교차' 기능을 통해 버퍼와 겹치는 카페를 골라내었다.

다음으로 네이버 지도 검색을 통해 각 카페의 아메리카노 가격을, 공정거래위원회 가맹사업거래 사이트의 정보공개서 검색을 통해 프랜차이즈 여부를 조사하였다. 이를 통해 고가 프랜차이즈, 중가 프랜차이즈, 저가 프랜차이즈 카페를 개별 시트화하였다.

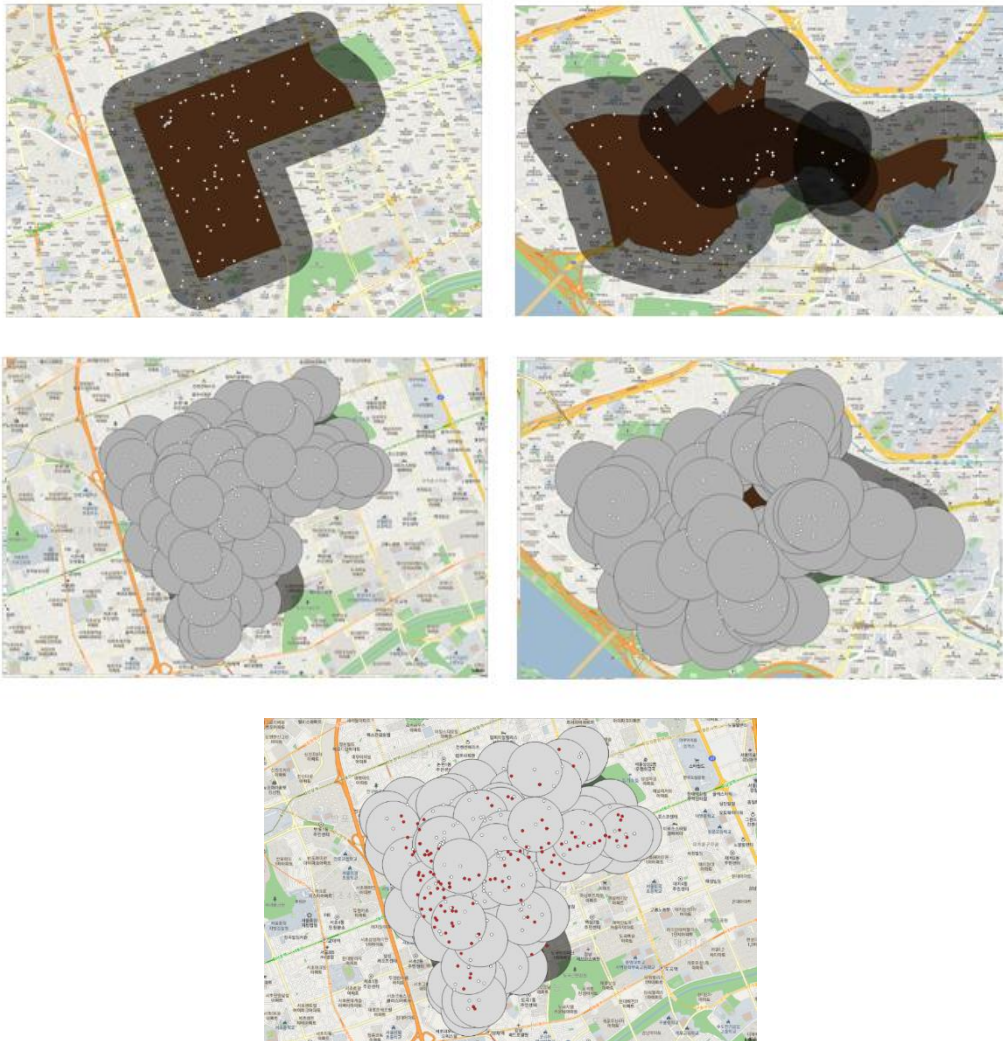
가격 기준은 여러 기사에 따라 아메리카노 가격이 4000 원 이상인 경우

고가 프랜차이즈, 2000 원 이상 4000 원 미만인 경우 중가 프랜차이즈, 2000 원 미만인 경우 저가 프랜차이즈로 분류하였다.

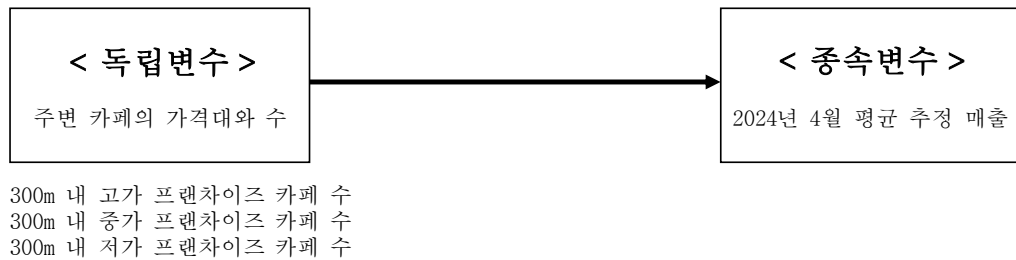
마지막으로 주변 카페 수 카운트를 위해 연구 대상 카페 데이터를 기준으로 300M 버퍼를 생성하였다. 앞서 만든 세 개의 시트를 불러온 후, ‘폴리곤에 포함하는 포인트 개수 계산’ 기능을 통해 각 시트에서 버퍼와 겹치는 카페의 개수를 세어 ‘300M 내 고가 프랜차이즈 카페 수’, ‘300M 내 중가 프랜차이즈 카페 수’, ‘300M 내 저가 프랜차이즈 카페 수’ 칼럼을 형성함으로써 독립변수를 생성하였다.

두 상권에서 수집한 데이터의 수는 도합 고가 프랜차이즈 224 개, 중가 프랜차이즈 189 개, 저가 프랜차이즈 197 개 이다.

[그림 2] QGIS 활용 분석을 통한 독립 변수 획득 과정



[그림 3] 회귀 분석 연구 모형



2. 연구 결과

2.1 오피스 상권 기초통계와 상관관계

오피스 상권 기초 통계 확인 결과 종속 변수가 넓은 분포를 나타내었으나, 모델의 성능이 충분하다고 판단되어 로그로 변환한 값을 사용하지는 않았다.

본 연구의 종속변수와 독립변수 사이에는 최소 0.12 ~ 최대 0.30 의 약한 양의 상관관계가 나타났다. 독립변수 간에는 분석에 영향을 미칠 만큼의 상관관계는 나타나지 않는 것으로 보인다.

<표 1> 오피스 상권 기초통계량

변수	Mean	S.D	1	2	3	4
제곱 미터 당 평균 추정 매출	232775.85	301683.649				
300M 내 고가 프랜차이즈 수	8.40	6.350	.304			
300M 내 중가 프랜차이즈 수	5.79	4.068	.124	.620	.	
300M 내 저가 프랜차이즈 수	4.99	4.079	.202	.708	.702	

2.2 오피스 상권 회귀 분석

다중회귀분석(Multiple linear regression analysis)이란 두 개 이상의 연속형 독립변수가 연속형 종속변수에 미치는 영향을 검증하는 방법이다. 독립변수를 가로축에, 종속변수를 세로축에 놓고 산점도를 그렸다고 가정했을 때, 산점도의 각 점인 ‘실제값’과 회귀분석을 통해 도출한 직선 위의 점인 ‘예측값’ 사이의 차이를 ‘잔차’라고 한다. 잔차제곱합을 최소화하는 직선을 3차원 이상의 공간에서 도출하는 분석 방법이 다중회귀분석이다.

300M 내 고가 프랜차이즈 카페의 수, 300M 내 중가 프랜차이즈 카페의 수, 300M 내 저가 프랜차이즈 카페의 수가 제곱 미터 당 추정 매출에 미치는 영향을 검증하기 위해, 다중회귀분석을 실시하였다.

그 결과 회귀모형은 통계적으로 유의하게 나타났으며($F=4.066$, $p=.009$), 회귀모형의 설명력은 10%(수정된 R 제곱은 7.5%)로 나타났다

한편 Durbin-Watson 통계량은 1.882로 잔차의 독립성 가정에는 문제가 없는 것으로 평가되었고, 분산팽창지수(Variance Inflation Factor: VIF)도 모두 10 미만으로 작게 나타나 다중공산성 문제 또한 없는 것으로 판단하였다.

회귀계수의 유의성 검증 결과, 300M 내 고가 프랜차이즈 카페 수($\beta=.353$, $p=.009$)가 정(+)적으로 유의하게 나타났다. 즉 오피스 300M 내 고가 프랜차이즈 카페 수가 많아질수록 제곱 미터 당 평균 추정 매출이 높아지는 것으로 나타났다.

<표 2> 오피스 상권 회귀 분석 결과

종속변수	독립변수	비표준화 계수		표준화 계수	t	p	VIF
		B	SE	β			
제공 미터 당 평균 추정 매출	(상수)	130173.922	49863.673		2.611	.010	
	300M 내 고 가 프랜차 이즈 수	16754.039	6276.766	.353	2.669**	.009	2.133
	300M 내 중 가 프랜차 이즈 수	-8955.629	9714.252	-.121	-.922	.359	2.097
	300M 내 저 가 프랜차 이즈 수	2752.060	10763.916	.037	.256	.799	2.589
$F(p)$		4.066(.009)					
adj. R^2		.075					
Durbin-Watson		1.882					

*p<.05, **p<.01, ***p<.001

2.3 핫플레이스 상권 기초통계와 상관관계

기초통계 확인 결과 종속 변수가 넓은 분포를 나타내었으나, 모델의 성능이 충분하다고 판단되어 로그로 변환한 값을 사용하지는 않았다.

본 연구의 종속변수와 독립변수 사이에는 최소 $-.225$ 에서 최대 $.077$ 의 약한 상관관계가 나타났다. 구체적으로, $.077$ 의 상관계수는 두 변수 사이에 매우 약한 양의 상관관계를 나타내며, $-.225$ 의 상관계수는 약한 음의 상관관계를 의미한다. 또한, $-.046$ 의 상관계수는 두 변수 사이에 거의 상관관계가 없음을 보여준다.

독립변수 간에는 분석에 영향을 미칠 만큼의 상관관계는 나타나지 않는 것으로 보인다.

<표 3> 핫플레이스 상권 기초통계와 상관관계

변수	Mean	S.D	1	2	3	4
제공 미터 당 평균 추정 매출	195185.47	175841.897				
300M 내 고가 프랜차이즈 수	5.18	4.292	.077			
300M 내 중가 프랜차이즈 수	2.85	1.657	-.225	.451		
300M 내 저가 프랜차이즈 수	3.75	3.243	-.046	.422	.419	

2.4 핫플레이스 상권 회귀 분석

300M 내 고가 프랜차이즈 카페의 수, 300M 내 중가 프랜차이즈 카페의 수, 300M 내 저가 프랜차이즈 카페의 수가 제곱 미터 당 평균 추정 매출에 미치는 영향을 검증하기 위해, 다중회귀분석을 실시하였다.

그 결과 회귀모형은 통계적으로 유의하게 나타났으며($F=3.373$, $p=.021$), 회귀모형의 설명력은 9%(수정된 R 제곱은 6.4%)로 나타났다

한편 Durbin-Watson 통계량은 1.712로 잔차의 독립성 가정에는 문제가 없는 것으로 평가되었고, 분산팽창지수(Variance Inflation Factor: VIF)도 모두 10 미만으로 작게 나타나 다중공산성 문제 또한 없는 것으로 판단하였다.

회귀계수의 유의성 검증 결과, 300M 내 중가 프랜차이즈 카페 수($\beta=-.325$, $p=.004$)가 부(-)적으로 유의하게 나타났다. 즉 300M 내 중가 프랜차이즈 카페 수가 많아질수록 제곱 미터 당 평균 추정 매출이 낮아지는 것으로 나타났다.

또한 300M 내 고가 프랜차이즈 카페 수($\beta=.226$, $p=.044$)가 정(+)적으로 유의하게 나타났다. 즉 오피스 상권에서와 마찬가지로 핫플레이스 상권에서도 300M 내 고가 프랜차이즈 카페 수가 많아질수록 제곱 미터 당 평균 추정 매출이 높아지는 것으로 나타났다.

표준화 계수의 절대값을 비교하면, 300M 내 중가 프랜차이즈 카페 수($\beta=-.325$), 300M 내 고가 프랜차이즈 카페 수($\beta=.226$) 순으로 제곱 미터 당 평균 추정 매출에 영향을 미치는 것으로 검증되었다.

<표 4> 핫플레이스 상권 회귀 분석 결과

종속변수	독립변수	비표준화 계수		표준화 계수	t	p	VIF
		B	SE	β			
제공 미터 당 평균 추정 매출	(상수)	246561.982	34613.362		7.123	.010	
	300M 내 고 가 프랜차 이즈 수	9272.094	4547.917	.226	2.039*	.044	1.369
	300M 내 중 가 프랜차 이즈 수	-34524.210	11760.170	-.325	-2.936**	.004	1.365
	300M 내 저 가 프랜차 이즈 수	-293.945	5914.606	.005	-.050	.960	1.322
$F(p)$		3.373(.021)					
adj. R^2		.064					
Durbin-Watson		1.712					

IV. 연구 2: 매출 상위 카페 리뷰 분석

1. 연구 방법

1.1. 연구 대상

본 연구의 대상은 앞선 연구의 기준 데이터가 된 카페 목록 중 제품 미터당 평균 추정 매출 상위 20 개 카페를 대상으로 하였다.

1.2. 자료 수집

본 연구의 리뷰 데이터는 기준 데이터 각 카페의 네이버 지도 리뷰 탭에서 수집하였다. 2022 년 4 월의 리뷰부터 수집하였는데, 이는 외식산업전망지수를 기준으로 코로나 19 라는 특수 상황에서의 리뷰 데이터를 배제하기 위한 것이다.

외식산업경기전망지수(KRBI)란, 매출액, 고객 수, 종업원 수 등 외식산업의 성장 및 위축 정도에 따라 종사자들의 판단, 예측의 변화 추이를 측정하여 산업경기동향을 분기별로 보여주는 체감지표이다. 즉, 외식업 종사자들이 체감하는 다음 분기 외식업의 전망을 의미한다.

코로나 19 발생 이전 시기인 2019 년 80 점대를 웃돌던 외식산업경기전망지수는 2020 년 코로나 19 의 영향으로 50 점 대까지 추락했다.

2021년 70 점대를 맴돌다가 2022년 2분기 90 점대로 급격한 회복세를 보였다. 이같은 2022년 2분기 외식 산업 경기 회복은 2022년 4월 18일 사회적 거리두기가 완전 해제되며 일어난 것으로 해석된다. 이에 따라 2022년 4월의 리뷰부터 수집함으로써 코로나 19가 리뷰에 미친 영향을 배제하고자 하였다.

[그림 4] 외식산업경기동향지수 추이 (조선비즈, 2022)



최종 수집된 리뷰의 개수는 다음과 같다.

<표 5> 상권 별 수집된 리뷰 수

	오피스 상권	핫플레이스 상권
리뷰 개수	3232 개	3766 개

1.3. 분석 절차

데이터 수집 과정은 다음과 같다. 먼저 기준 데이터를 제공 미터 당 추정 매출 평균을 기준으로 내림차순으로 정렬하여 제공 미터 당 추정 매출 평균 상위 20 개 카페를 추려내었다. 다음으로, 각 카페를 네이버 지도에 검색하였고, 자동화 웹스크래핑 도구 리스틀리(listly) 프로그램을 사용하여 각 카페의 2022 년 4 월부터 2024 년 3 월까지의 방문자 리뷰 데이터를 수집하였다.

데이터 전처리 과정은 다음과 같다. 수집된 리뷰를 파이썬 KoNLPy 라이브러리의 OpenKoreanText(Okt)를 활용하여 전처리하였다. 먼저 “가성비”, “휘낭시에”, “아인슈페너”, “밀크티”, “디카페인”, “웨이팅”, “키오스크”, “재방문”, “걸바속촉” 등 분석되진 않았으나 리뷰에 자주 활용된 단어를 분석에 추가하였다. “이쁘다” - “예쁘다”, “티라미슈” - “티라미수”, “주인” - “사장”, “케익” - “케익” - “케이크” 등 의미가 동일함에도 다르게 표기된 단어들은 분석의 정확도를 올리기 위해 하나의 단어로 통일하였다. 모든 군집에서 가장 큰 비율로 나타난 “맛있다”, “좋다”의 경우, 빈도가 너무 커서 분석에 방해되는 단어에 해당하였다. 보다 특징적인 의미 도출을 위해 제거하였다.

다음으로 띄어쓰기 단위로 텍스트를 토큰화한 후, 어간을 추출하였고 품사 태깅에서는 명사, 동사, 형용사만을 활용하였다.

분석 방법으로는 키워드 빈도 분석과 TF-IDF, LDA 토픽모델링을 활용하였다. 키워드 빈도 분석은 키워드의 단순 노출 순위와 정도를 보여주며, TF-IDF 는 단어의 빈도와 역 문서 빈도를 사용하여 단어들마다 중요한 정도에

따라서 가중치를 부여하는 방법이다. 본 연구에서는 LDA 토픽모델링을 통해 리뷰를 군집화하였다. LDA 토픽모델링은 여러 주제가 혼합된 문서에서 텍스트의 모든 단어로부터 잠재 주제를 추론하는 기법이다. 두 상권 모두 선행 연구에 따라 분석 결과의 복잡도(perplexity)가 낮고 주제 내 단어 간 응집도(coherence)가 낮게 나타난 3을 기준으로 적정 토픽의 수를 정하였고, 결과 해석 시에는 차별되는 단어를 찾아 토픽별로 의미를 구분하고 명칭을 부여하였다.

분석의 틀로는 카페 서비스품질 측정 선행연구(엄성원,2022)에서 제시된 서비스 품질 매트릭스에 의한 키워드 추출 예시를 활용하였다. 연구에 제시된 샘플키워드에 더해 관련성이 존재한다고 판단되는 키워드들을 차원에 따라 해석하였다.

유형성은 눈으로 확인 가능한 물리적인 시설로서, ‘인테리어’, ‘가구’, ‘위치’, ‘비주얼’, ‘예쁘다’ 및 메뉴와 관련한 키워드를 이에 포함하였다. 신뢰성은 서비스와 서비스 제공자에 대한 신뢰를 의미하며, ‘가격’, ‘가성비’, ‘품질’, ‘위생’, ‘맛’, ‘원두’, ‘산미’ 등을 이에 포함하였다. 대응성은 즉각적인 응대를 의미하며 ‘빠르다’, ‘예약’, ‘친절’, ‘만족’ 등을 이에 포함하였다. 확신성은 직원의 전문성 및 신뢰감을 의미하며, ‘수상’, ‘전문가’, ‘최고’, ‘믿다’ 등의 키워드를 이에 포함하였다. 마지막으로 공감성은 고객에 대한 개별적인 관심을 의미한다. 이에 따라 고객 맞춤형 서비스와 관련된 키워드인 ‘감동’, ‘친절’, ‘감사’ 등을 포함하였다.

2. 연구 결과

2.1. 오피스 상권 TF 및 TF-IDF 분석

오피스 상권 추정 매출 평균 상위 20 개 비프랜차이즈 카페의 리뷰에 대한 빈도분석을 실시한 결과, 커피(1012)가 가장 높은 빈도를 나타냈으며, 그 다음으로 스콘(438)이 가장 높은 빈도를 나타냈다. 오피스 상권의 주요 고객층인 직장인들은 카페에 오래 머무르기보다는 주로 테이크 아웃을 통한 구매를 선호한다. 이에 따라 인테리어 등의 카페 내부 요소보다는 메뉴의 맛을 가장 중요시함을 알 수 있다. 이와 유사한 의미의 라떼(287), 맛집(174), 디저트(166) 역시 상위 10 위권 내에 형성하였다. 또한 친절(330), 사장(127), 직원(80) 등 서비스 제공자와 관련된 단어들이 상위 30 위권 내를 형성하였다. 자주(144), 근처(106), 회사(77) 등은 일회성 방문 고객이 아닌 카페 주변에서 생활 권역을 가지는 회사원 고객들의 높은 재방문을 시사한다. 가성비(103)는 23 위로 나타났다. 합리적인 가격에 대한 믿음을 의미하는 가성비는 많은 고객을 유인하는 매력적인 요소이다. 25 위인 주문(90) 키워드를 통해 픽업에 대한 높은 니즈를 확인하였다. 비교적 시간 여유가 적은 회사원들은 출근 시간이나 점심 식사 후에 네이버 주문 등의 서비스를 통해 메뉴를 미리 주문하여 픽업하는 시스템에 대한 높은 만족도를 보이고 있었다.

단어의 문서 내 중요도를 수치로 보여주는 TF-IDF 분석 결과, 가성비(23 위 → 6 위)와 저렴하다(19 위로 새로 랭킹)의 순위가 크게

상승되었다. 이는 오피스 상권에서 가격 경쟁력의 중요성을 시사한다. 또한 ‘항상’의 순위가 40위에서 26위로 크게 상승되었다. 이는 오피스 상권 고객의 일상적이고 반복적인 구매를 시사한다.

<표 6> 오피스 상권 TF 분석 결과

	단어	빈도(개)		단어	빈도(개)
1	커피	1012	21	예쁘다	107
2	스콘	438	22	근처	106
3	먹다	349	23	가성비	103
4	친절	330	24	여기	91
5	라떼	287	25	주문	90
6	카페	185	26	고소하다	86
7	오다	185	27	종류	85
8	가다	180	28	가격	85
9	맛집	174	29	직원	80
10	디저트	166	30	밀크티	80
11	보다	157	31	않다	78
12	많다	155	32	회사	77
13	마시다	147	33	다양하다	74
14	자주	144	34	원두	73
15	사장	127	35	비엔나	73
16	분위기	119	36	크림	72
17	티라미수	114	37	없다	70
18	방문	112	38	아메리카노	68
19	매장	110	39	휘낭시에	68
20	음료	109	40	향상	67

<표 7> 오피스 상권 TF-IDF 분석 결과

	단어	TF-IDF		단어	TF-IDF
1	커피	292.3853679	21	보다	31.18686505
2	친절	110.7814821	22	매장	30.66944309
3	스콘	92.61962296	23	고소하다	29.27299496
4	라떼	85.79165401	24	방문	28.51335397
5	먹다	67.92211354	25	밀크티	28.44430268
6	가성비	57.74862027	26	항상	28.25619776
7	맛집	53.31278669	27	가격	27.8067319
8	디저트	53.24155104	28	근처	26.89020335
9	자주	47.25068248	29	여기	25.70926224
10	가다	44.20364211	30	직원	25.61591648
11	분위기	44.07065768	31	종류	24.79347805
12	오다	43.19619601	32	회사	23.64937704
13	마시다	42.09713628	33	아메리카노	23.35234031
14	카페	40.64856865	34	휘낭시에	23.15580016
15	많다	37.05663077	35	마들렌	22.10517348
16	음료	36.26563577	36	주문	21.12466242
17	예쁘다	33.14015459	37	최고	20.86790168
18	티라미수	33.07432148	38	깔끔하다	20.59949833
19	저렴하다	32.93372319	39	다양하다	20.58636685
20	사장	32.90967548	40	비엔나	20.57688313

2.2. 핫플레이스 상권 단어빈도(TF) 및 TF-IDF 분석

핫플레이스 상권 추정 매출 평균 상위 20 개 비프랜차이즈 카페의 리뷰에 대한 빈도분석을 실시한 결과, 먹다(1257)가 가장 높은 빈도를 나타냈으며, 커피(945), 수플레(761) 순으로 높은 빈도를 나타냈다. 이는 음식의 맛 자체를 즐기기 위해 핫플레이스 비프랜차이즈 카페에 방문하는 고객들이 다수 존재함을 시사한다.

오피스 상권과의 키워드 차이를 중심으로 살펴보면, 먼저 디저트의 종류에서 차이가 나타났다. 스콘, 티라미수 등 주로 아메리카노에 곁들여 향유되는 비교적 가벼운 디저트들이 나타난 오피스 상권 리뷰와 달리, 핫플레이스 상권 리뷰에서는 수플레(761), 케이크(464), 팬케이크(186)처럼 객단가나 체류시간을 높이는 디저트들이 나타났다. 혹은 다쿠아즈(246), 휘낭시에(232)처럼 베이커리 전문 카페에서 ‘구움과자’로 불리며 높은 인기를 구가하는 메뉴들이 나타났다. 메뉴를 묘사하는 키워드도 다르게 나타났다. 오피스 상권에서는 ‘고소하다’, ‘다양하다’가 나타난 반면, 핫플레이스 상권에서는 ‘부드럽다’, ‘달달’이 나타나며 보다 달콤한 메뉴에 대한 선호를 확인할 수 있었다.

또한 홍대(328)를 통해 ‘홍대 카페를 찾아보다가 방문하였다’는 다수의 리뷰를 확인할 수 있었다. 즉 핫플레이스 방문 고객들이 핫플레이스를 특별한 장소로서 인식하며 친목 도모나 데이트 등의 목적으로 까다롭게 카페를 선택하여 방문하는 경향이 있음을 시사한다.

장소를 묘사하는 키워드로 오피스 상권에는 없던 인테리어(213), 넓다(175)가 등장하였고, TF-IDF 분석에서 순위가 상승하며 카페 인테리어와

크기의 중요도를 입증하였다. 이는 체류시간이 긴 핫플레이스 카페의 특성상 카페의 인테리어와 크기에 민감한 고객들이 존재함을 시사한다.

마지막으로 TF-IDF 분석 결과 ‘귀엽다’가 상위 키워드로 새롭게 등장하였다. ‘예쁘다’는 오피스 상권에서도 상위 키워드로 나타났지만, ‘귀엽다’는 핫플레이스 상권의 상위 키워드로만 확인할 수 있었으며, 인테리어나 매장 콘셉트, 플레이팅, 캐릭터를 활용한 디저트의 비주얼 등에서 귀여움을 느끼는 고객 리뷰들을 통해 핫플레이스 상권의 비프랜차이즈 카페에서 소비자가 귀여움을 느낄만한 요소들을 제공하는 것의 중요성을 확인할 수 있었다.

<표 8> 핫플레이스 상권 TF 분석 결과

	단어	빈도(개)		단어	빈도(개)
1	먹다	1257	21	홍대	328
2	커피	945	22	푸딩	318
3	수플레	761	23	들다	297
4	카페	727	24	크다	288
5	디저트	701	25	크림	285
6	친절	670	26	라떼	269
7	오다	601	27	없다	246
8	음료	503	28	다쿠아즈	246
9	딸기	494	29	메뉴	244
10	분위기	473	30	휘낭시에	232
11	케이크	464	31	다음	220
12	사장	447	32	인테리어	213
13	예쁘다	438	33	부드럽다	191
14	보다	431	34	추천	186
15	맛집	413	35	팬케이크	186
16	방문	393	36	달달	183
17	가다	388	37	복숭아	180
18	넘다	381	38	되다	176
19	많다	365	39	넓다	175
20	매장	334	40	마시다	175

<표 9> 핫플레이스 상권 TF-IDF 분석 결과

	단어	TF-IDF		단어	TF-IDF
1	커피	208.6258744	21	많다	63.30932165
2	친절	168.7945514	22	휘낭시에	61.50217228
3	먹다	153.3653247	23	보다	60.91328819
4	디저트	135.4815806	24	크다	58.21280133
5	수플레	125.7264886	25	홍대	56.69002735
6	분위기	110.9805712	26	들다	49.65278964
7	카페	108.4031622	27	인테리어	49.07891428
8	음료	98.70872953	28	크림	48.0417433
9	케이크	97.45433416	29	메뉴	46.60311855
10	오다	94.3900292	30	넓다	43.89165302
11	맛집	91.88286721	31	최고	42.65638528
12	예쁘다	85.69859549	32	직원	42.5077676
13	딸기	85.10795413	33	다음	42.17204919
14	사장	83.84014969	34	팬케이크	41.92334536
15	넘다	78.24601809	35	마시다	41.82138351
16	가다	70.70841138	36	복숭아	39.6184085
17	방문	69.65980341	37	없다	37.4717674
18	다쿠아즈	64.70814347	38	자주	37.13468697
19	라떼	64.48358063	39	부드럽다	36.47862668
20	매장	64.43968799	40	귀엽다	35.12893223

2.3. 워드 클라우드

[그림 5] 오피스 상권 워드 클라우드



2.4. 토픽 모델링 결과

오피스 상권의 토픽들을 살펴보면 토픽 1은 ‘커피’가 가장 높은 비율로 나타났으며 ‘라떼, 가다, 자주, 근처, 가성비, 마들렌, 사장, 주문’이 차별 단어로 나타났다. 이를 통해 전반적인 큰 토픽의 주제는 커피 구매를 위해 일상적으로 방문하는 고객군과 관련되어 있다. 서비스 품질 매트릭스의 관점에서 살펴보면 ‘자주’, ‘근처’는 커피 구매를 위해 근처 생활권에서 일상적으로 방문하는 고객의 수요를 시사하며 매장 위치 관련 키워드로 유형성에 해당한다. ‘가성비’는 가격에 대한 신뢰성을 의미한다. ‘마들렌’의 경우 한 카페에서 다회차 방문 고객을 대상으로 서비스되는 품목임을 확인하였다. 이는 비프랜차이즈 카페 로열티 프로그램에 대한 고객의 만족과 그 중요성을 보여준다는 점에서 공감성에 해당한다. ‘주문’은 대응성에 해당한다. 시간 여유가 적은 직장인들을 위해 오피스 상권에서는 ‘주문’ 시스템을 활용한 빠른 픽업 서비스의 중요성을 확인할 수 있다.

다음으로 토픽 2은 ‘친절’이 가장 높은 비율로 나타났으며 ‘저렴하다, 항상, 가격, 맛집, 분위기, 감사하다, 직원, 매일, 디저트, 세트’와 같은 차별 단어가 나타났다. 이를 통해 전반적인 큰 토픽의 주제는 커피 및 디저트 구매를 위해 일상적으로 방문하는 고객군과 관련되어 있음을 알 수 있다. ‘매일’은 근처 생활권에서 일상적으로 방문하는 고객의 수요를 시사하며 매장 위치 관련 키워드로 유형성에 해당한다. 서비스 품질 매트릭스의 관점에서 살펴보면 ‘항상’, ‘맛집’은 일관된 높은 품질에 대한 신뢰성을 의미한다. ‘저렴하다’, ‘가격’은 가격 경쟁력의 중요성을 시사하며 서비스의 낮은 가격에 대한 신뢰성을 보여준다. ‘감사하다’는 서비스에 대한 고객의 만족과 그 중요성을 보여준다는 점에서 공감성에 해당한다.

‘세트’를 통해 리뷰에서 식사 대용으로 커피와 샌드위치 세트를 구매하는 고객들의 높은 수요와 만족도를 확인하였다. 이는 오피스 상권 고객 맞춤형 메뉴라는 점에서 공감성에 해당한다.

마지막으로 토픽 3은 ‘커피’가 상위 키워드로, ‘스콘, 많다, 보다, 카페, 오다, 디저트, 예쁘다, 방문, 종류, 맛집, 티라미수’가 차별 단어로 나타났다. 이를 통해 전반적인 큰 토픽의 주제는 카페 탐방을 위해 비일상적으로 방문하는 고객군과 관련되어 있음을 알 수 있다. 서비스 품질 매트릭스의 관점에서 살펴보면 ‘스콘, 예쁘다, 티라미수, 디저트, 종류’는 유형성에 해당한다. ‘예쁘다’는 인테리어 혹은 메뉴 외형의 중요성을 시사한다. ‘디저트’, ‘종류’는 다양한 디저트 메뉴 제공의 필요성을, ‘스콘’, ‘티라미수’는 인기 메뉴 존재의 중요성을 시사한다. ‘맛집’은 카페를 맛집으로 인식하는 고객의 신뢰성을 의미한다.

이처럼 오피스 상권의 토픽 1에서는 커피 구매를 위해 일상적으로 방문하는 고객군의 리뷰에서 유형성, 신뢰성, 공감성, 대응성을, 토픽 2에서는 디저트 구매를 위해 일상적으로 방문하는 고객군의 리뷰에서 유형성, 공감성을, 토픽 3에서는 카페 탐방을 위해 비일상적으로 방문하는 고객군의 리뷰에서 유형성, 신뢰성을 확인할 수 있었다.

다음으로 핫플레이스 상권의 토픽들을 살펴보면 토픽 1은 ‘떡다’가 상위 키워드로, ‘오다, 딸기, 케이크, 보다, 크림, 다음’이 차별 단어로 나타났다. 이를 통해 전반적인 큰 토픽의 주제는 디저트 구매를 위해 방문하는 고객군과 관련되어 있음을 알 수 있다. 서비스 품질 매트릭스의 관점에서 살펴보면 ‘딸기, 케이크, 크림’은 유형성에 해당하며, 맛이 좋고 다양한 메뉴 제공의 필요성을 시사한다. ‘다음, 오다’를 통해 카페 이용 경험에 만족한 고객들이 핫플레이스 상권 비프랜차이즈 카페 재방문을 원하는 리뷰를

확인할 수 있었다. 이는 고객 만족에 대한 리뷰로서 공감성에 해당한다.

핫플레이스 상권 리뷰에서 토픽 2 은 ‘친절’ 이 가장 높은 비율로 나타났으며 ‘친절, 사장, 커피, 매장, 자주’ 등의 차별 단어가 나타났다.

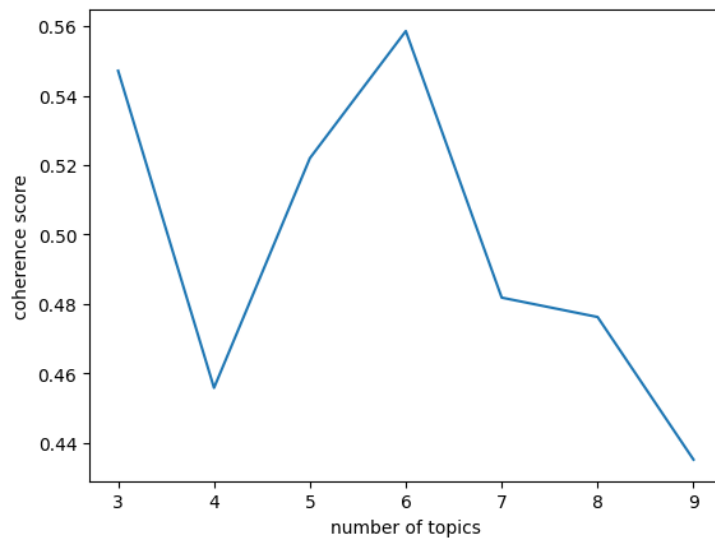
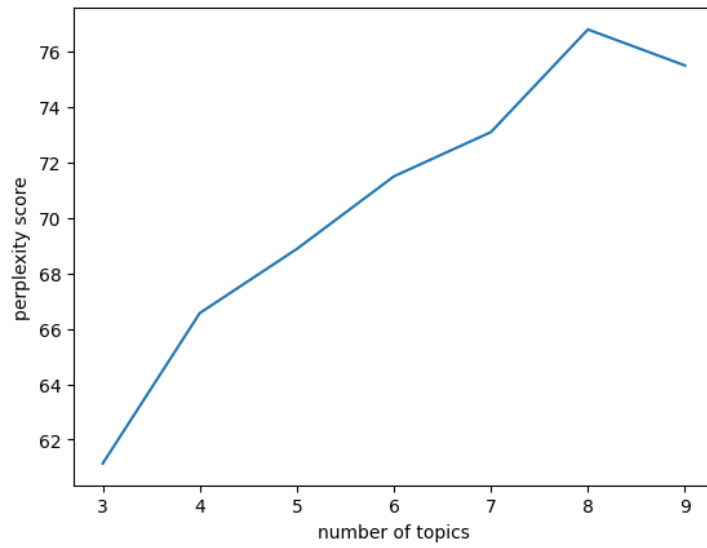
‘커피’의 경우 타 군집에서도 나타났으나 타 군집 대비 가장 높은 순위를 보여 차별 단어로 선정하였다. 이를 통해 전반적인 큰 토픽의 주제는 커피 구매를 위해 일상적으로 방문하는 고객군과 관련되어 있음을 알 수 있다. 서비스 품질 매트릭스의 관점에서 살펴보면 ‘자주’는 근처 생활권에서 일상적으로 방문하는 고객의 수요를 시사하며 매장 위치와 관련되어 유형성을 보여준다. 핫플레이스 상권 리뷰에서 전반적으로 비일상적 방문이 주를 이루었지만, 직장인의 리뷰에서는 일상적 방문을 확인할 수 있었다. ‘친절’은 핫플레이스 상권에서 커피를 마시는 것을 목적으로 카페 방문 시 친절함 서비스의 중요하게 고려됨을 시사하며 대응성에 해당한다.

마지막으로 토픽 3 은 ‘수플레’가 상위 키워드로, ‘수플레, 예쁘다, 맛집, 브런치, 팬케이크, 넓다, 많다, 홍대’가 차별 단어로 나타났다. 이를 통해 전반적인 큰 토픽의 주제는 카페 탐방을 위해 방문하는 고객군과 관련되어 있음을 알 수 있다. 서비스 품질 매트릭스의 관점에서 살펴보면 ‘수플레, 예쁘다, 팬케이크’는 유형성에 해당하며, 비주얼이 아름다운 디저트 메뉴 제공의 필요성을 시사한다. 또다른 유형성 키워드인 ‘넓다, 많다’를 통해 카페의 다양한 메뉴와 큰 매장 크기에 만족한 고객들의 리뷰를 확인할 수 있었다. ‘홍대’는 홍대입구역과 가까운 카페의 위치에 대한 만족을 나타냈으며 이 역시 유형성에 해당한다. ‘맛집’은 주로 만족한 메뉴의 이름과 하나로 묶여 나타났으며, 맛에 대한 신뢰성을 보여주었다.

이처럼 핫플레이스 상권의 토픽 1에서는 디저트 구매를 위해 방문하는 고객군의 리뷰에서 유형성과 공감성을, 토픽 2에서는 커피 구매를 위해

방문하는 고객군의 리뷰에서 유형성과 대응성을, 토픽 3에서는 카페 탐방을 위해 방문하는 고객군의 리뷰에서 유형성과 신뢰성을 확인할 수 있었다.

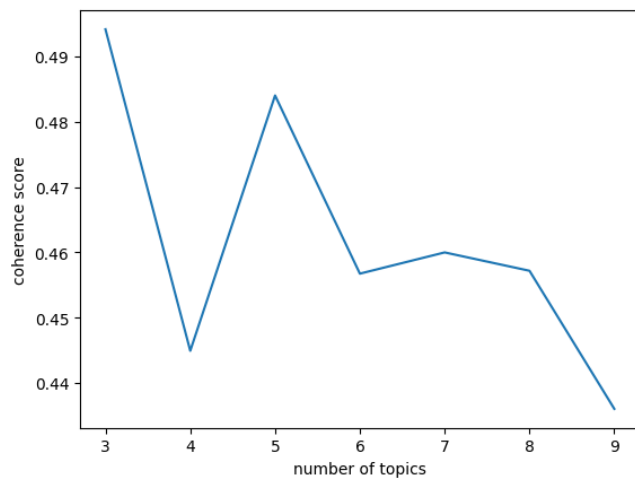
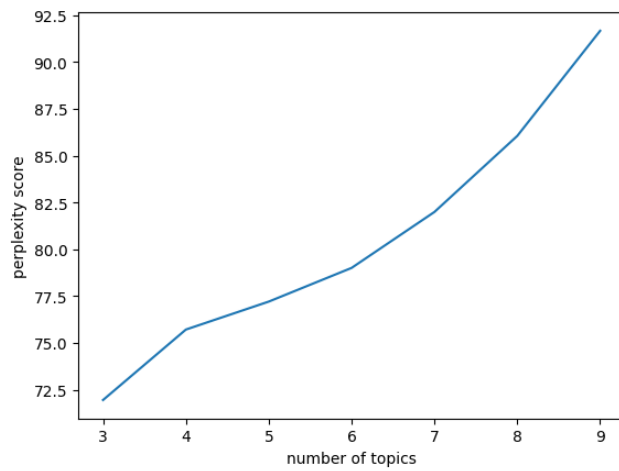
[그림 7] 오피스 상권 perplexity & coherence



<표 10> 오피스 상권 토픽 모델링 결과

	TOPIC 1		TOPIC 2		TOPIC 3	
순위	키워드	가중치	키워드	가중치	키워드	가중치
1	커피	0.027	친절	0.023	커피	0.031
2	라떼	0.019	커피	0.017	스콘	0.023
3	먹다	0.015	저렴하다	0.015	먹다	0.016
4	마시다	0.009	항상	0.015	많다	0.015
5	가다	0.008	가격	0.013	보다	0.014
6	자주	0.008	마시다	0.010	카페	0.014
7	친절	0.007	맛집	0.008	오다	0.013
8	근처	0.007	분위기	0.008	친절	0.011
9	가성비	0.006	음료	0.008	디저트	0.011
10	마들렌	0.005	출근	0.007	예쁘다	0.010
11	고소하다	0.005	감사하다	0.007	가다	0.010
12	오다	0.005	직원	0.007	종류	0.009
13	주문	0.005	매일	0.007	맛집	0.009
14	사장	0.005	디저트	0.007	티라미수	0.009
15	보다	0.005	세트	0.006	방문	0.008
차별 단어	라떼, 가다, 자주, 근처, 가성비, 마들렌, 고소하다, 주문, 사장		저렴하다, 항상, 가격, 맛집, 분위기, 감사하다, 직원, 매일, 디저트, 세트		스콘, 많다, 보다, 카페, 오다, 디저트, 예쁘다, 방문, 종류, 맛집, 티라미수	

<그림 8> 핫플레이스 상권 perplexity & coherence



<표 11> 핫플레이스 상권 토픽 모델링 결과

	TOPIC 1		TOPIC 2		TOPIC 3	
순위	키워드	가중치	키워드	가중치	키워드	가중치
1	먹다	0.028	친절	0.023	수플레	0.031
2	오다	0.012	푸딩	0.017	예쁘다	0.023
3	딸기	0.011	사장	0.015	맛집	0.016
4	카페	0.011	커피	0.015	디저트	0.015
5	디저트	0.010	먹다	0.013	브런치	0.014
6	케이크	0.010	카페	0.010	커피	0.014
7	커피	0.009	많다	0.008	친절	0.013
8	수플레	0.009	디저트	0.008	넓다	0.011
9	보다	0.009	음료	0.008	카페	0.011
10	가다	0.008	자주	0.007	팬케이크	0.010
11	방문	0.007	가다	0.007	많다	0.010
12	크림	0.007	매장	0.007	음료	0.009
13	음료	0.006	보다	0.007	분위기	0.009
14	다음	0.005	분위기	0.007	홍대	0.009
15	맛집	0.009	오다	0.006	케이크	0.008
차별 단어	오다, 딸기, 케이크, 보다, 크림, 다음		친절, 푸딩, 사장, 커피, 매장, 자주		예쁘다, 맛집, 브런치, 팬케이크, 넓다, 많다, 홍대	

V. 추가 분석 1: 핫플레이스 중가 프랜차이즈 리뷰 분석

1. 연구 방법

1.1. 연구 대상

본 연구 대상은 서교동 경계 기준으로 300M 내에 있는 중가 프랜차이즈 카페들이다. 연구 1 에서 수집한 핫플레이스 중가 프랜차이즈 카페 리스트로부터 서교동 행정동 경계 300M 내의 카페를 QGIS 를 활용하여 추려내었다.

1.2. 자료 수집

첫 번째로 오픈업에서 각 카페의 추정매출을 조사하여 인허가데이터의 소재지 면적 칼럼으로 나누었고, 제곱 미터 당 추정 매출 칼럼을 생성하였다. 이를 내림차순 정렬하여 상위 10 개 카페를 선정하였다.

두 번째로 본 연구의 리뷰 데이터는 기준 데이터 각 카페의 네이버 지도 리뷰 탭에서 수집하였다. 연구 2 와 마찬가지로 2022 년 4 월의 리뷰부터 수집하였다. 최종 수집된 리뷰의 개수는 3518 개이다.

데이터 분석 과정은 연구 2 와 같다.

2. 연구 결과

2.1 TF 및 Tf-idf 분석 결과

핫플레이스 상권 추정 매출 평균 상위 20 개 중가 프랜차이즈 카페의 리뷰에 대한 빈도분석을 실시한 결과, 망고(373)가 가장 높은 빈도를 나타냈으며, 먹다(307), 친절(268) 순으로 높은 빈도를 나타냈다. 망고(373), 도넛(112), 빙수(109), 코코넛(48) 등의 키워드를 통해 커피 판매가 증가되는 일반적인 카페보다는 특징적인 메뉴의 프랜차이즈들이 핫플레이스 상권에서 선전하고 있음을 확인할 수 있었다.

특히 가성비(107), 가격(61), 저렴(41)과 같이 가격에 대한 언급이 빈도 순위 상위를 차지하였다. 이를 통해 핫플레이스 중가 프랜차이즈 카페의 소비자는 저렴한 메뉴를 기대하며 카페를 방문함을 확인할 수 있다. 또한 넓다(84), 시원(43)과 같이 장소의 특징에 대한 키워드는 장소의 편의성에 대한 기대로 방문하는 고객의 수요를 시사한다.

빈도 분석의 역시(39 위에서 24 위)와 TF-IDF 분석 결과 순위가 상승한 항상(38 위에서 27 위)과, 새로 나타난 자주(30 위)는 일관된 품질과 서비스를 제공하는 프랜차이즈의 특성을 선호하는 다수의 고객을 시사한다.

이외에도 핫플레이스 상권의 상위 매출 프랜차이즈 카페들을 조사하는 과정에서 매달 이뤄지는 통신사 할인에 대한 높은 선호를 확인할 수 있었다.

<표 12> 핫플레이스 상권 중가 프랜차이즈 TF 분석 결과

	단어	빈도(개)		단어	빈도(개)
1	망고	373	21	자다	61
2	먹다	307	22	밀크	60
3	친절	268	23	식후땀	59
4	있다	218	24	주문	58
5	많다	187	25	메뉴	58
6	매장	146	26	최고	58
7	음료	134	27	들다	57
8	커피	123	28	오다	54
9	사람	115	29	되다	52
10	도넛	112	30	홍대	51
11	빙수	109	31	딸기	50
12	가성비	107	32	아니다	49
13	없다	95	33	코코넛	48
14	가다	87	34	그렇다	47
15	넓다	84	35	방문	46
16	직원	83	36	크다	43
17	보다	79	37	시원	43
18	나오다	78	38	항상	42
19	마시다	65	39	역시	42
20	가격	61	40	저렴	41

<표 13> 핫플레이스 상권 중가 프랜차이즈 카페 TF-IDF

	단어	TF-IDF		단어	TF-IDF
1	친절	143.5322	21	메뉴	21.90599
2	망고	91.03542	22	마시다	21.71969
3	먹다	74.7645	23	자다	20.57675
4	많다	53.72205	24	역시	19.54639
5	있다	53.712	25	가격	19.43786
6	가성비	47.97812	26	보다	18.44793
7	매장	44.10604	27	항상	18.29944
8	음료	43.87023	28	시원	17.59953
9	커피	42.06372	29	저렴	17.53218
10	도넛	41.85475	30	자주	17.22471
11	넓다	37.07471	31	딸기	15.77542
12	사람	34.57478	32	오다	15.50472
13	빙수	34.43089	33	자리	15.15809
14	밀크	25.9554	34	홍대	14.96498
15	나오다	25.64809	35	방문	14.86924
16	최고	24.95343	36	들다	14.50771
17	가다	24.67359	37	도너츠	14.05327
18	직원	24.5366	38	코코넛	13.72056
19	없다	23.44095	39	들르다	13.54956
20	식후땀	21.94333	40	추천	13.23065

[그림 9] 핫플레이스 상권 증가 프랜차이즈 워드 클라우드



VI. 추가 분석 2: 작성 회차에 따른 오피스 상권 리뷰 비교

1. 연구 방법

1.1 연구 대상

본 연구는 앞선 연구 2와 마찬가지로 제품 미터 당 추정 매출 상위 20개의 오피스 상권 비프랜차이즈 카페를 대상으로 하였다.

1.2 자료 수집

본 연구의 리뷰 데이터는 연구 2에서 수집한 오피스 상권 리뷰를 다시 활용하였다. 총 3232개의 리뷰를 리뷰 작성 회차에 따라 분류하였다. 최종적으로 분류된 리뷰의 개수는 다음과 같으며, 분석 절차는 연구 2와 동일하되, 품사 태깅에서는 명사만을 활용하였다.

<표 14> 작성 회차 별 수집된 리뷰 수

	1 회차 리뷰	다회차 리뷰
리뷰 개수	1786 개	1445 개

2. 연구 결과

2.1 리뷰 작성 회차에 따른 TF 순위 비교

곳, 좋아요 등 무의미한 리뷰를 제외하자 다회차 리뷰의 개수가 1회차 리뷰에 비해 현저히 감소하였다. 이에 따라 빈도 분석 비교는 빈도가 아닌 순위를 기준으로 시행하였다. 오피스 상권 추정 매출 평균 상위 20 개 비프랜차이즈 카페의 1회차 리뷰와 다회차 리뷰에 대한 빈도분석을 실시한 결과, 다음과 같은 특징이 나타났다.

첫 번째로, 1회차 리뷰에서는 분위기(7위), 인테리어(23위)가 높은 순위를 기록한 반면 다회차 리뷰에서는 40위권 안에서 두 단어를 확인할 수 없었다. 이는 오피스 상권에서 카페에 오래 머무르려는 목적보다는 주로 테이크아웃 등을 통한 제품 구매를 목적으로 재방문이 이뤄짐을 시사한다.

1회차 리뷰보다 다회차 리뷰에서 높게 나타난 단어들은 다음과 같다. 먼저 다회차 리뷰에서 가성비(4위)가 높은 순위로 나타났다. 이는 단골 유치를 위해 가격 경쟁력 확보가 중요함을 시사한다. 최고(19위), 제일(38위)과 같은 단어들은 타 카페와의 비교 우위를 시사하며, 자주(3위), 항상(9위)과 같은 키워드들은 동일한 카페를 재방문하고 리뷰까지 재작성한 고객의 높은 방문 빈도를 시사한다. 오늘(7위)과 아침(10위)은 일상 생활 속 재방문을 보여주며, 회사(14위), 출근(21위), 출근길(29위)은 높은 출근길 픽업 수요를 시사한다. 주문(6위), 네이버(35위)는 네이버 주문을 활용한 고객들의 높은 만족도를 보여준다.

<표 15> 리뷰 작성 회차에 따른 오피스 상권 TF 분석 결과

	1 회차 리뷰		다회차 리뷰	
	단어	빈도(개)	단어	빈도(개)
1	커피	683	커피	329
2	스콘	415	라떼	89
3	라떼	198	자주	58
4	카페	159	가성비	51
5	디저트	138	마들렌	41
6	맛집	138	주문	38
7	분위기	112	오늘	37
8	사장	109	맛집	36
9	매장	94	항상	34
10	방문	92	여기	33
11	음료	89	디저트	28
12	자주	86	아침	27
13	티라미수	83	아메리카노	27
14	종류	75	회사	26
15	크림	70	카페	26
16	비엔나	68	스콘	23
17	직원	67	가격	22
18	밀크	67	원두	22
19	근처	65	최고	21
20	가격	63	방문	20

	1 회차 리뷰		다회차 리뷰	
	단어	빈도(개)	단어	빈도(개)
21	휘낭시에	59	출근	20
22	여기	58	친절	18
23	인테리어	58	사장	18
24	아메리카노	54	바닐라	17
25	선물	54	메뉴	17
26	가성비	52	매일	17
27	주문	52	근처	16
28	회사	51	이용	16
29	원두	51	출근길	16
30	추천	46	매장	16
31	사람	45	세트	15
32	최고	44	밀크	14
33	다음	44	직원	13
34	포장	41	디카페인	13
35	달달	38	네이버	13
36	점심	38	샌드위치	12
37	산미	36	딸기	11
38	메뉴	34	제일	11
39	항상	33	산미	11
40	아인슈페너	33	종류	10

VII. 결론

1. 연구 결과 요약

비프랜차이즈 주변 카페의 가격대와 수가 추정매출에 영향을 미치는지 알아본 연구 1의 결과는 다음과 같다. 오피스 상권과 핫플레이스 상권 모두에서 고가 프랜차이즈 카페 수가 많을수록 추정매출이 증가하는 양(+)의 관계를 보였지만, 핫플레이스 상권에서는 특별히 중가 프랜차이즈 수가 증가할수록 추정매출이 감소하는 모습을 보였다. 이를 통해 고가 프랜차이즈 카페는 두 상권 모두에서 상생 관계에 있다는 것을 확인하였다. 이는 고가 프랜차이즈 카페의 입지 특성이 주로 지하철 역과 가까이 위치해 있어 유동인구가 많고, 많은 수요로 이어져 추정 매출에 긍정적인 영향을 주었음을 시사한다. 또한 핫플레이스 상권에서 중가 프랜차이즈 카페 수가 증가할수록 추정매출이 감소한 연구 결과에 따라, 핫플레이스 상권에서 중가 프랜차이즈 카페가 비프랜차이즈 카페와 경쟁 관계에 있음을 확인하였다.

상권 별 리뷰 특성 차이를 알아본 연구 2의 결과는 다음과 같다. 빈도 분석 및 TF-IDF 분석 결과, 두 상권 간 디저트 종류와 메뉴의 맛을 묘사하는 단어에서 차이가 나타났다. LDA 토픽모델링 결과 두 상권 모두에서 고객 리뷰는 세 가지 군집으로 나뉘었으며, 전체적으로 커피 구매, 디저트 구매, 그리고 카페 탐방이라는 비슷한 방문 목적을 도출하며 군집화되었다. 이를 통해 두 상권에서 유사한 방문 목적을 가진 고객들이 방문하고 있음을 확인할 수 있었다.

한편, 유사한 군집이 형성되었지만 각 상권의 특성에 따라 소비자가

집중하는 서비스 요소는 다르게 나타났다. 오피스 상권 경우 주로 직장인의 일상적 방문이 이루어진 반면, 핫플레이스 상권의 경우 주로 비일상적 방문이 이루어졌고 커피 구매 군집에서만 일부 직장인의 일상적 방문이 나타났다. 서비스 품질 매트릭스의 관점에서 살펴본 결과 토픽에 따라 각기 다른 다양한 특성을 보여주었다.

오피스 상권에서는 토픽 1에서 커피 구매를 위해 일상적으로 방문하는 고객군의 리뷰에서 유형성, 신뢰성, 공감성, 대응성을, 토픽 2에서 디저트 구매를 위해 일상적으로 방문하는 고객군의 리뷰에서 유형성, 공감성을, 토픽 3에서 카페 탐방을 위해 비일상적으로 방문하는 고객군의 리뷰에서 유형성, 신뢰성을 확인할 수 있었다.

핫플레이스 상권에서는 토픽 1에서 디저트 구매를 위해 방문하는 고객군의 리뷰에서 유형성과 공감성을, 토픽 2에서 커피 구매를 위해 방문하는 고객군의 리뷰에서 유형성과 대응성을, 토픽 3에서 카페 탐방을 위해 방문하는 고객군의 리뷰에서 유형성과 신뢰성을 확인할 수 있었다

2. 학문적 시사점

본 연구의 학문적 시사점은 다음과 같다. 본 연구는 빅데이터 기반 상권 분석 플랫폼 오픈업을 통해 연구 1의 종속 변수와 연구 2의 연구 대상 선정 기준으로 개별 매장의 추정 매출 평균 데이터를 활용하였다. 이를 통해 지역별 데이터, 상권별 데이터를 통한 분석이 주로 이루어지던 기존의 연구들과 달리 차별화된 연구가 가능했다. 즉 마케팅 측면에서 개별 매장의 추정 매출에 대한 정량적인 평가가 이루어졌다. 이를 통해 독립 변수가 개별 매장의 매출에 미치는 영향을 알 수 있었고 상위 매출 카페를 선정할 수 있었다.

연구 1에서는 주변 300M 내 프랜차이즈 카페가 추정 매출에 미치는 영향을 분석함으로써 상권별 주변 프랜차이즈 카페와의 상생 및 경쟁 관계를 규명하였다.

연구 2에서는 코로나 19 이후 작성된 국내 카페 이용자 온라인 리뷰를 분석함으로써 카페에 대한 최신의 인식을 확인하였다. 또한 카페 이용자의 관점에서 방문 목적 관련 인식은 커피 구매, 디저트 구매, 카페 탐방으로 구성됨을 규명하였다. 또한 SERVQUAL 모형을 적용한 리뷰 분석을 통해 SERVQUAL 모형을 카페 연구 분야에서 활용할 수 있음을 재검증하였다.

3. 실무적 시사점

본 연구의 실무적 시사점은 다음과 같다. 우선 연구 1에서 핫플레이스 카페에서 주변 300M 내에 중가 프랜차이즈 카페 수가 많을수록 추정 매출의 감소가 나타났다. 이를 통해 핫플레이스 상권에서의 비프랜차이즈 카페 보호를 위한 프랜차이즈 출점 제한 제도 도입의 필요성을 입증하였다. 즉 프랜차이즈 출점 제한 제도를 단계적으로 도입하고자 할 때, 핫플레이스 상권에 우선적으로 도입을 고려할 필요성을 확인하였다.

또한 중가 프랜차이즈와 경쟁 관계를 형성하고 있는 핫플레이스 상권에서는 경쟁력 강화를 위한 사업주의 심사숙고가 요구됨을 확인하였다. 연구 2의 핫플레이스 중가 프랜차이즈의 여러 리뷰에서 멤버십 적립 및 통신사 할인 행사에 대한 긍정적 반응을 확인하였다. 즉 핫플레이스 비프랜차이즈 카페는 이를 벤치마킹하여 쿠폰 제도 등을 활용한 로열티 프로그램, 정기 할인 이벤트 등을 시도하거나, 프랜차이즈처럼 제품과 서비스의 일관된 품질을 확보하는 등의 노력이 요구된다. 이를 통해 경쟁력을 강화할 수 있음을 시사한다.

연구 2에서는 토픽 모델링을 통해 커피 중심, 디저트 중심, 카페 탐방 중심의 세 고객 군집을 발견하였다. 분석 결과 나타난 상권 특성을 반영하여 각각의 군집을 타겟팅하는 포지셔닝 전략을 구현할 수 있을 것이다.

추가 분석에서는 오피스 상권의 높은 다회차 리뷰 수가 확인됨에 따라 오피스 상권 고객 단골 고객 유치 전략 수립의 중요성을 발견하였다.

‘출근길’, ‘세트’ 등의 키워드 발견을 통해 출근길의 직장인을 타겟으로 하는 식사 대용 세트 메뉴의 필요성이 확인되었다.

4. 한계점 및 향후 연구 방향

본 연구의 한계점은 다음과 같다. 첫 번째로 연구 1에서 분석 대상 데이터의 양이 많지 않다는 것이다. 그 이유로는 먼저 오픈업에서 분석 대상인 비프랜차이즈 카페의 추정 매출 데이터가 검색되지 않는 경우 분석에서 제외되었다는 점이 있다. 또다른 이유로는 네이버 지도에 ‘역삼 1동 카페’, ‘서교동 카페’를 검색한 결과 인허가 데이터에서 발견하지 못한 카페들이 다수 발견되었다는 점이 있다. 검색 결과 데이터를 연구 데이터에 포함했으나, 검색 결과에 포함되지 않은 카페들도 다수 존재하리라 예상된다.

이와 관련한 두 번째 한계점은 연구 1에서 행정동 경계 기준 주변 600M 카페 데이터 세트를 형성할 때에도 전체 카페를 조사하기 어려웠다는 점이다. 즉, 연구 데이터에 전체 카페가 포함되지 못했다는 한계를 지닌다.

향후 연구 방향은 다음과 같다. 소비자의 카페 인지 단계에서 정보 습득이 일어나는 또다른 플랫폼에 대한 추가적인 연구가 필요하다. 특별히 인스타그램에 주목할 필요가 있다. 데이터 습득 과정에서 다수의 핫플레이스 카페에서 인스타그램 계정을 운영하고 있음을 확인할 수 있었기 때문이다. 이에 따라 향후 카페 매출 영향 요인을 연구할 때 인스타그램 계정 유무, 인스타그램 카페 태그 횟수, 인스타그램에 나타난 카페 정보의 양 등을 고려한다면 유의미한 시사점을 발견할 수 있으리라 기대한다.

VIII. 참고문헌

- 최누리 and 김인신. (2022). 부산시 카페 생존분석: 프랜차이즈와 독립카페 간 비교. 호텔경영학연구, 31(7), 105-118.
- 민철기 and 강창덕. (2021). 상권의 공간적 확산에 따른 상업시설 생존율과 생존요인 비교 - 홍대지역 음식점을 중심으로 -. 서울도시연구, 22(2), 17-38.
- 김성아, 김병준, 박수현, 김규림 and 김홍순. (2023). 서울시 핫플레이스의 입지 특성 분석: 2017 년과 2022 년의 비교를 중심으로. 한국도시지리학회지, 26(3), 33-49.
- 신재욱. (2014). 프랜차이즈 커피 브랜드 차별화 전략 방안에 관한 연구. 조형미디어학, 17(1), 105-116.
- 배준호, 이준재 and 박정숙. (2014). 커피전문점의 선택속성에 대한 세분시장별 만족도에 관한 연구: 강릉 커피거리를 중심으로. 관광레저연구, 26(1), 375-390.
- 최성임, 임은순 and 문혜선. (2012). 선호커피유형에 따른 세분시장의 특성. 한국콘텐츠학회 논문지, 12(6), 475-485.
- 임병학 and 엄성원. (2022). 카페 서비스품질 측정에 관한 연구:마케팅 애널리틱스 접근법. 경영교육연구, 37(5), 1-21.
- 김용원 and 김용규. (2021). 동별 음식점 매출과 블로그 게시글 수 사이의 상호 관계 연구 : 서울특별시를 중심으로. 외식경영연구, 24(2), 123-146.

윤동화 and 안진현. (2022). 군집분석을 이용한 제주도 지역별 음식점 특성 분석: 네이버 플레이스 키워드 리뷰 중심으로. 한국산학기술학회논문지, 23(12), 173-179.

정동영, 김태경 and 김지영. (2022). 고객의 방문빈도가 입소문 효과에 미치는 영향 : 소비자리뷰의 감정강도를 중심으로. 대한경영학회지, 35(3), 1-26.

정의석, 김병석 and 황조혜. (2020). 언어네트워크 분석을 활용한 핫플레이스 이미지 연구 -서울특별시 지역별 맛집을 중심으로-. 관광학연구, 44(6), 9-31.

신우진 and 문소연. (2011). 프랜차이즈 커피전문점의 입지특성이 매출액에 미치는 영향 분석. 부동산학연구, 17(2), 111-123.

안예린, 이현희, 유성민, 김서연 and 박민서. (2023). 머신러닝 알고리즘을 이용한 커피 프랜차이즈 실패 요인 분석에 관한 연구. 대한산업공학회지, 49(1), 37-45.

유예림. (2023). 옆집 또 카페... "출점 제한" 외치는 점주들. *머니투데이*. <https://news.mt.co.kr/mtview.php?no=2023050122524477275#:~:text=%EA%B3%B5%EC%A0%95%EA%B1%B0%EB%9E%98%EC%9C%84%EC%9B%90%ED%9A%8C%EB%8A%94%202012,%EC%A7%80%EC%A0%81%EC%9D%B4%20%EB%82%98%EC%99%94%EA%B8%B0%20%EB%95%8C%EB%AC%B8%EC%9D%B4%EB%8B%A4>”

양범수. (2022). 사회적 거리두기 해제로 외식경기 회복... “인플레로 인한 소비위축은 악영향. *조선비즈*. <https://biz.chosun.com/distribution/food/2022/07/15/2ZLSJSXLD5AHNPYY7LG>

T4NEVA/

통계청(2022). 「2022 서비스업 조사」.

https://kostat.go.kr/statDesc.es?act=view&mid=a10501010000&sttr_cd=S009003

와이즈앱(2024). 「모빌리티 앱 사용자 수 3700 만 명 시대! 모빌리티 앱 이용 현황」.

<https://www.wiseapp.co.kr/insight/detail/501>

한국데이터산업진흥원 데이터안심구역(2020). 「지역 현황 - 지역별 유동인구 현황 분석」.

https://dsz.kdata.or.kr/svc/page/application/data_analysis_vitalize.do

서울특별시(2017). 「서울생활문화자료조사 - 홍대앞」.

https://museum.seoul.go.kr/archive/archiveNew/NR_archiveList.do?ctgryId=CTGRY684&type=A

ABSTRACT

Factors affecting Independent Cafés Sales Using Platform Data

Hye-Min Song

Department of Business Administration

The Graduate School

Sookmyung Women's University

The saturation of the café market has led to a significant increase in café closures. Independent cafés, in particular, face dual challenges from the price competitiveness of low-cost franchise cafés and the financial power of high-end franchise cafés. This study aims to provide insights into the introduction of franchise café entry restrictions and the development of positioning strategies for independent cafés through two exploratory studies focused on independent cafés in hotplace and office districts. To achieve this, estimated sales data were collected from the big data commercial district analysis platform, OpenUp, and review data were gathered from Naver Maps. In Study 1, the impact of the price and number of surrounding cafés on estimated sales was examined through multiple regression analysis. The results showed that in both districts, the presence of more high-end franchise cafés was associated with an increase in average estimated sales per square meter, while in hotplace districts, an increase in the number of mid-range franchise cafés was associated with a decrease in average estimated sales per square meter. Study 2 used LDA topic modeling on reviews from the top 20 cafés in terms of average estimated sales per square meter to identify differences in the characteristics of the two districts. The results revealed that independent cafés in both office and hotplace districts could be categorized

into three clusters based on the purposes of coffee purchase, dessert purchase, and café exploration, with specific differences in content reflecting the characteristics of each district.

Key words : independent cafés, online reviews, topic modeling, text mining