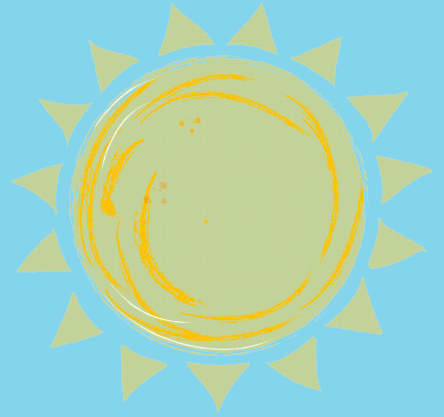


빅데이터

기말 프로젝트 보고서



20172543 우정호

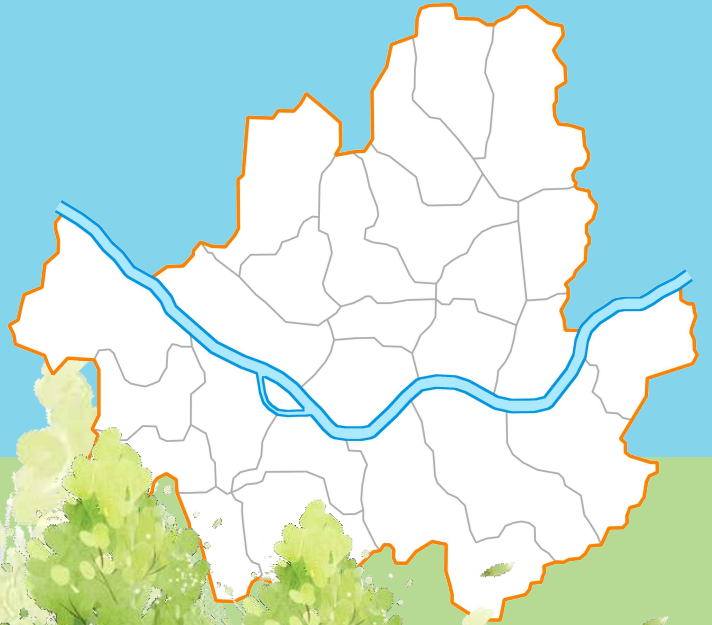
상권 개요

키워드 분석

상권 분석

인구 현황

상권 요약



명동 이야기



June, 2022



백석대학교

학교 백석대학교

담당 이시은 교수님

학부 컴퓨터공학부

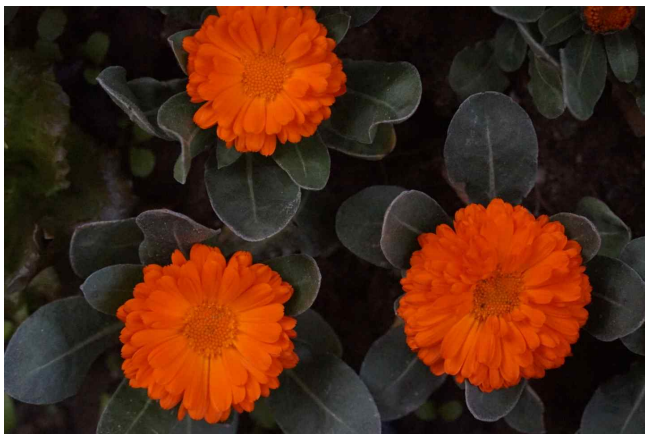
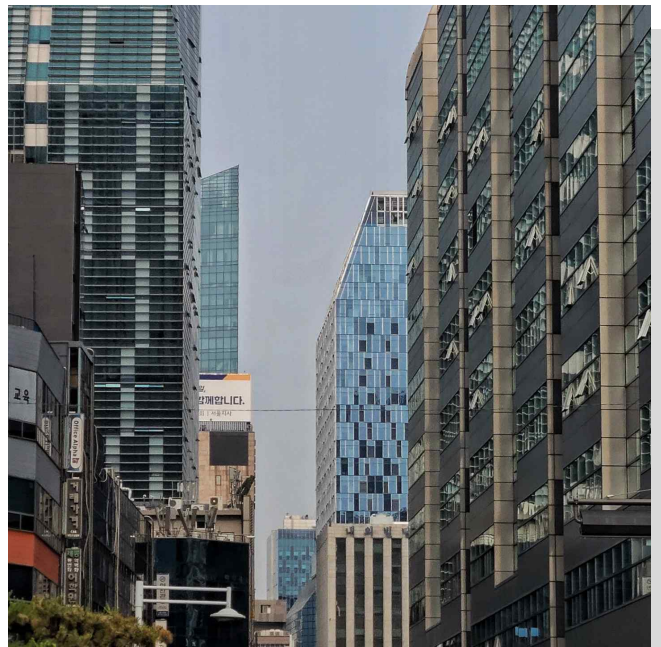
영역 빅데이터

학과 멀티미디어학과

주제 상권분석

목차

- 04 (1) 상권 개요
 - 상권 선정
 - 상권의 지리적 위치
 - 상권의 특징
- 06 (2) 키워드 분석
 - 워드 클라우드 분석
 - 검색 분석
- 09 (3) 상권 분석
 - 업종 분석
 - 매출 분석
 - 상권주도 업종
- 12 (4) 인구 현황
 - 유동인구
 - 상주인구
- 14 (5) 상권 요약
 - 상권 분석 정리
 - 가설 확인
 - 가설 확인 참고자료
 - 우수점포 소개
- 18 (6) 마무리
 - 개선사항 및 소감
- 19 (7) 기타
 - 코드 및 설명
 - 참고 및 활용 자료
 - 전체 코드



상권 개요

1.1) 상권 선정

기말 프로젝트를 진행하면서 조사 할 상권을 어떻게 선정 할지 고민하였다. 처음에는 내가 거주하고 있는 지역의 상권을 조사하거나 혹은 경기도민이 많이 접할 수 없는 지방의 상권을 조사하는 방향으로 프로젝트를 계획했으나 반 학기 동안 강의를 들으며 가장 많이 다루어 본 데이터를 가지고 보다 심도 있게 분석을 하고 싶다는 생각을 하게 되었다. 내가 프로젝트를 진행하며 조사할 상권은 서울특별시를 대상으로 하며, 그 이유는 앞서 말한 바와 같이 개인 과제와 팀 과제를 수행하며 가장 많이 접할 수 있었던 서울특별시의 상권을 통해 더 높은 수준의 결과물을 내보이기 위함이다.

서울특별시 내에도 많은 상권이 있으니 그 중에서 하나의 상권을 특정해야한다. 상권은 매출의 증감을 시각화해보고 특이점이 있는 상권 탐색하는 것으로 프로젝트를 진행한다. 상권 선정은 최근 데이터는 코로나19가 상권의 매출에 유의미한 영향을 끼쳤을 것으로 예상되기에 이를 위해 코로나19가 발병하기 이전인 2019년의 데이터와 가장 최근 데이터인 2021년 데이터를 비교 분석하여 매출이 큰 증감이 일어난 상권을 색출한다. 상권을 특정하기 위해 상권 매출의 연 단위 합계액을 기준으로 하였다.

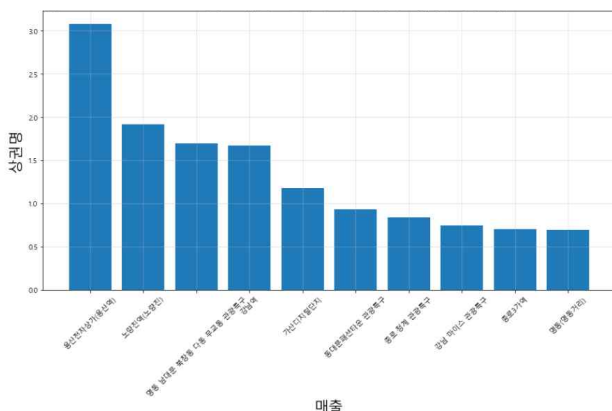
서울 열린 데이터 광장에서 제공하는 ‘서울시 우리마을가게 상권분석 서비스(상권-추정매출)’ 자료를 통해 서울시 2019년과 2021년의 서울 시내 상권의 연 단위 매출 합계를 10순위까지 비교하였다.

두 자료 간 비교를 통해 확인 할 수 있었던 것은 2019년 상위 10개 상권 가운데 8개의 상권이 2021년에도 상위 10위를 기록한 것을 확인 할 수 있다. 해당

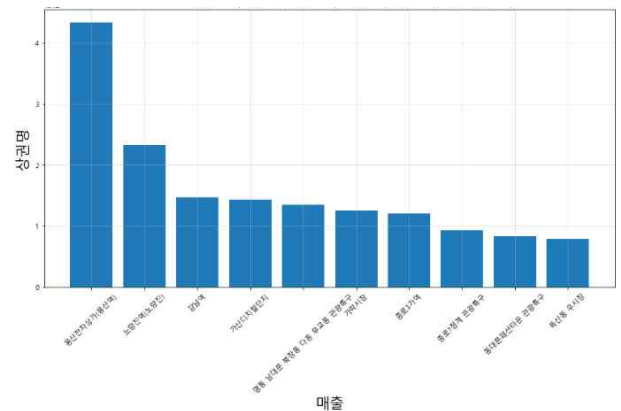
	2019년	2021년
1	용산전자상가	용산전자상가(-)
2	노량진역	노량진역(-)
3	명동관광특구	강남역(1▲)
4	강남역	가산디지털단지(1▲)
5	가산디지털단지	명동관광특구(2▼)
6	동대문관광특구	가락시장(13▲)
7	종로·청계 관광특구	종로3가역(2▲)
8	강남 마이스 관광특구	종로·청계 관광특구(1▼)
9	종로3가역	동대문 관광특구(3▼)
10	명동	독산동 우시장(49▲)

표[1-1-1]

상권은 용산전자상가, 노량진역, 강남역, 가산디지털단지, 명동관광특구, 종로3가역, 종로·청계 관광특구 그리고 동대문 관광특구이며 관광특구 3곳은 순위가 하락하였다. 용산전자상가와 노량진역은 2019년과 동일하게 각각 1위와 2위를 기록하였으며 순위가 상승한 상권은 한 계단 상승한 가산디지털단지와 강남역, 두 계단 상승한 종로3가역 그리고13계단 상승해 6위에 위치한 가락시장 상권과 무려 49계단 상승해 10위에 위치한 독산동 우시장 상권이다. 기존 2019년에 10위권이었지만 순위권 밖으로 밀려난 상권은 강남 마이스 관광특구와 명동이 유일하며 강남 마이스 상권은 8위에서 14위, 명동은 10위에서 49위로 하락하였다. 전체적으로 관광특구의 매출은 하락하였고 역세권 상권은 매출이 상승했으며 가락시장과 독산동 우시장 등 농산물/도축 등의 도소매를 취급하는 상권이 크게 매출이 상승했다고 분석 할 수 있다. 다음 상권 가운데 코로나의 장기화로 도소매업이 다시 성장세를 이루는 것과 관광객 봉쇄로 인한 관광특구의 순위 하락, 질병에 상대적으로 영향을 받지 않는 전자상가의 매출 유지는 사실 당연한 수순이라 볼 수 있다. 상권 가운데 명동은 순위가 39계단 하락한 것은 물론이고 매출은 19년의 46% 수준에 그쳤다. 이는 하나의 특이점이라 판단 할 수 있으므로 본 프로젝트는 명동(명동거리) 상권을 프로젝트 대상으로 특정하고 프로젝트를 진행하는 것으로 한다.



사진[1-1-1]



사진[1-1-2]

프로젝트를 진행할 상권을 선정하였는데, 다른 지표
를 고려하지 않고 단순히 매출액 합계로 순위를 계산
한 만큼 이는 정확한 지표가 아니며 코로나와의 직접
적인 관련성이 있다고 말 할 수 없음을 밝힌다. 앞선
과정은 상권을 선정하기 위한 과정이었을 뿐이며 다
만, 코로나19의 직간접적인 영향이 있었을 것이라 생
각되는 만큼 “매출액의 상승에 코로나19가 직간접적인
영향을 끼쳤을 것이다.”라는 가설을 두고 상권의 조사
분석 말미에 해당 가설의 사실 여부 또한 조사한 결
과를 삽입하도록 하겠다.

본 프로젝트의 목표는 최근 매출이 크게 증가한 상
권인 ‘명동(명동거리)상권’을 대상으로 ①해당 상권의
전반적인 분석[본 주제]과 ②매출액 증가와 코로나19
의 영향 유무 확인[소 주제]을 위함이다.

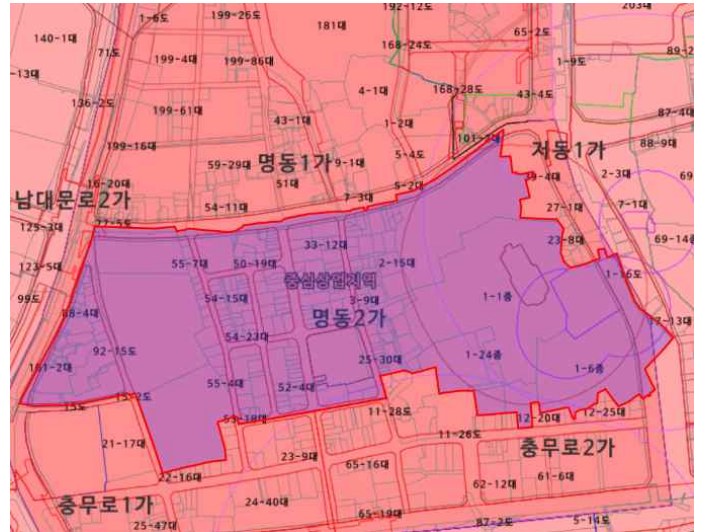
1.2) 상권의 지리적 위치

- 서울시 중구 명동에 위치한 명동(명동거리) 상권은
을지로입구역과 명동역 모두 도보 5분 이내로 걸
어갈 수 있는 조건을 가지고 있다.
- 명동(명동거리)상권은 국내외에서 유명한 상권으로
을지로입구역 5번 출구부터 명동역 6번 출구까지
이어진 상권이며 옆에 명동성당이 위치하고 있다.
- 명동은 행정구역 상 명동1·2가, 충무로1·2가, 을지
로1·2가 등을 포함하는 지역이지만 보편적으로 명
동거리는 명동1가와 명동2가를 일컫는 상권이며
두 지역을 합친 면적은 0.91km제곱이다.
- 명동 상권은 국토계획법에 따른 지역지구로는 중심
상업지역[도심의 상업 및 업무기능의 확충을 위하여 필요한
지역]에 해당하고 용도지구로는 방화지구[화재의 위험
을 예방하기 위한 지구]에 해당하며 용도구역으로는 제
1종지구단위계획구역[토지이용을 합리화, 구체화하고 도시
또는 농, 산, 어촌의 기능을 증진하며 미관을 개선하고 양호한
환경을 확보하기 위하여 수립한 구역]에 해당한다.



사진[1-2-1]. 명동거리 주변 입지

- 100여 개의 버스 노선과 지하철1·2·3·4호선이 지
나가는 교통의 요충지이다.
- 명동거리를 업종별로 구분하면 크게 중앙로, 1번가,
3번가, 명동길, 명동1가로 구분할 수 있는데 중앙
로는 은행·대형판매매장·화장품, 1번가는 쇼핑
물·먹거리, 3번가는 소매점·먹거리, 명동길은 명동
성당·화장품 그리고 명동1가는 판매점·식당이 주로
위치하고 있다.



사진[1-2-2]. 명동거리의 용도지역



사진[1-2-3]. 명동거리의 지역지구

1.3) 상권의 특징

- 우리나라에서 가장 많은 유동인구가 집계되는 지
역 중 한 곳으로 여성 유동인구 비율이 더 높다.
- 명동역 대로변에 위치한 숙박시설, 영화관, 명동
상권 내에 위치한 복합쇼핑센터, 학교와 주한중국
대사관까지 위치하고 있는 여러 유형의 복합 상권
으로 특성을 규정하기 어렵다.
- 평당 1억 9600만원으로 우리나라 최고의 공시지가
를 기록하고 있는 지역이다.

키워드 분석

키워드의 분석은 분석을 진행하는 과정에서 대상이 가지고 있는 특징을 직관적으로 알아볼 수 있기 때문에 명동 상권을 대상으로, 키워드의 분석을 진행한다. 키워드의 분석은 네이버 개발자에서 제공하는 네이버 API서비스와 네이버 광고에서 제공하는 키워드 검색 시스템을 이용해 각각 워드 클라우드의 분석과 검색어의 분석을 목표로 본 목차를 진행한다.

2.1) 워드 클라우드 분석

워드 클라우드의 분석에 앞서 출력을 하기 위해 워드 클라우드 출력에 사용될 검색 단어가 있어야하는데 나는 ‘명동’, ‘명동거리’ 그리고 ‘명동맛집’ 3개의 검색어를 가지고 프로젝트를 진행하며 분석의 다양성을 위해 3개의 단어를 뉴스와 블로그, 2개의 카테고리를 통해 출력한다. 그리고 다시 워드 클라우드에서 시각적으로 눈에 띄는 키워드를 추출해내 카테고리 별 공통 분모 및 차별점을 찾는 작업을 반복한다.



사진[2-1-1]. 뉴스에서 출력한 ‘명동’ 워드 클라우드



사진[2-1-2]. 블로그에서 출력한 ‘명동’ 워드 클라우드

첫 번째로, 단어 ‘명동’을 대상으로 워드 클라우드를 출력하였다. 뉴스에서는 ‘H온드림’과 ‘오프닝 임팩트’가, 블로그에서는 ‘정성본’과 ‘교자’의 키워드가 눈에 들어온다. ‘H온드림’은 현대차 그룹과 현대 정몽구 재단이 함께 추진하는 스타트업 프로젝트로 해당 이벤트의 ‘오프닝 임팩트’ 즉, 개최가 5월 30일 명동에서 이

루어졌기 때문에 뉴스에서 집중적으로 보도된 것을 확인할 수 있다. 반면 블로그에서의 ‘정성본’은 청담에 본점, 명동에 2호점 그리고 수도권을 대상으로 체인점을 운영하고 있는 샤브샤브 전문점인 정성본을 의미하며 ‘교자’는 명동에 위치하면서 ‘서울 미쉐린 가이드 2022’에 등재된 칼국수 및 만두 전문점인 교자를 의미한다. 뉴스와 블로그 양쪽 모두 현대그룹에서 운영하는 온드림 관련 내용이 키워드로 도출된 것을 공통 분모로 확인할 수 있다. 반면, 뉴스는 정책 및 사업과 관련한 키워드, 블로그는 요식업과 관련된 키워드가 주를 이루었다고 분석할 수 있다.



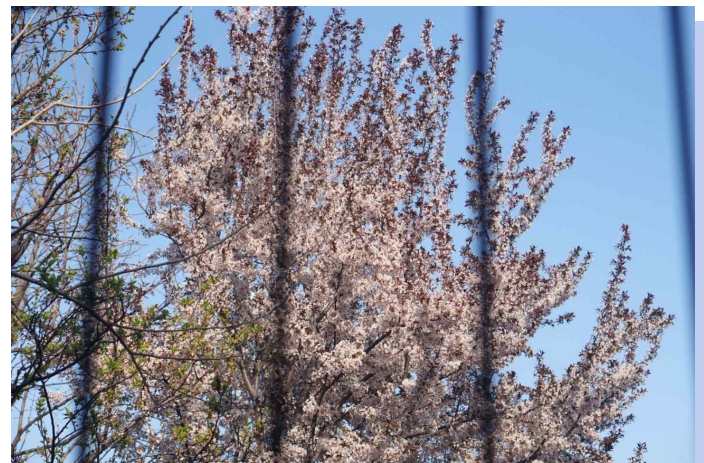
사진[2-1-3]. 뉴스에서 출력한 ‘명동맛집’ 워드 클라우드



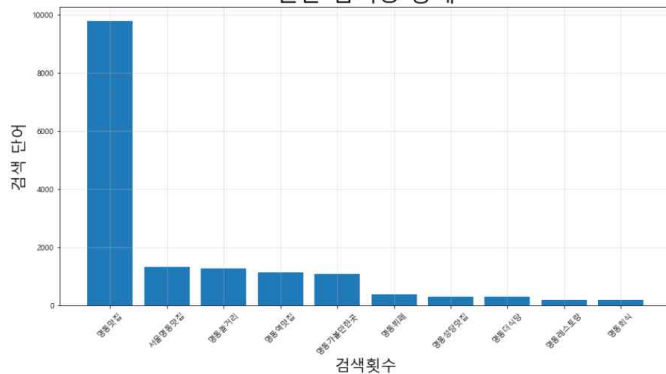
사진[2-1-4]. 블로그에서 출력한 ‘명동맛집’ 워드 클라우드

두 번째로, 단어 ‘명동맛집’을 대상으로 워드 클라우드를 출력하였다. 뉴스에서는 ‘약채락’과 ‘제천’이, 블로그에서는 ‘샤브샤브’와 ‘정성본’이 키워드인 것을 확인할 수 있다. 우선적으로 명동이라는 지명은 서울시에만 있는 지명이 아니며 충청북도 제천시에도 명동이라는 지명이 있다. 약채락이라는 단어는 충북 제천시의 한식전문점 약채락을 의미하며 같이 키워드로 언급된 ‘제천’ 역시 같은 이유로 볼 수 있다. 이 부분에서 알 수 있는 것은 상권의 규모 면에서 서울의 명동과 제천의 명동은 비교 대상이 아님에도 불구하고 키워드로 노출이 되었다는 이야기는 최근, 충북 제천시 그 중에서도 명동에의 관심이 증가하고 있다고 유추해 볼 수 있다.

를 기준으로 살펴본다면 ‘서울명동맛집’, ‘명동놀거리’, ‘명동역맛집’, ‘명동가볼만한곳’ 등 양쪽 모두 크게 키워드에 있어서 차이가 나는 부분이 있거나 순위에 변동이 있지 않다. 특이점이라고 한다면 ‘명동맛집’을 기준으로 검색한 연관 키워드이므로 ‘명동맛집’의 집계표본이 압도적으로 많다는 것과 ‘명동맛집’ 그리고 ‘서울명동맛집’의 경우 PC와 모바일의 월 평균 클릭 수가 크게 차이난다는 점이다. 다음으로는 데이터를 시각화한 자료를 보도록 하겠다.

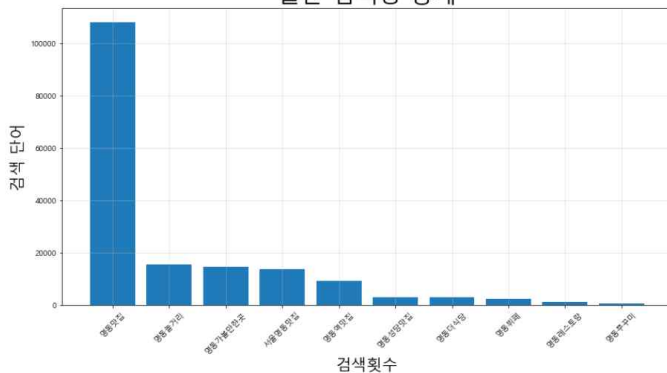


월간 검색량 통계



사진[2-2-3]. 월간 검색량 통계(PC)

월간 검색량 통계



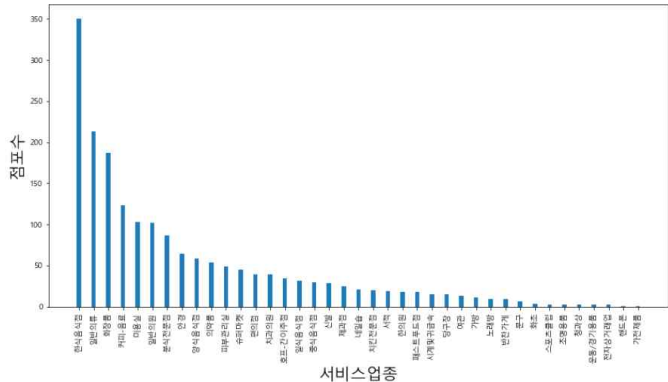
사진[2-2-4]. 월간 검색량 통계(모바일)

다음 그래프를 통해 확인할 수 있는 것은 PC와 모바일의 집계 데이터는 Y축의 주 단위 값이 각각 10,000과 100,000으로 10배 차이가 나는데 불구하고 각 키워드 항목별 검색 비율이 거의 동일하다는 점을 알 수 있는데, 검색 이용자가 PC와 모바일 양쪽이 거의 동일하다고 볼 수 있다.

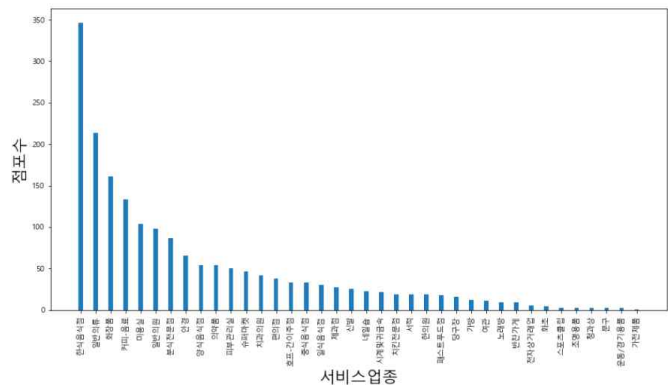
검색의 분석을 통해 알 수 있었던 것은 명동은 검색하는 고객층이 PC와 모바일 양쪽이 비슷한 특징을 가지고 있으며, 해당 고객들은 ‘맛집’, ‘놀거리’ 등 키워드에 직설적인 표현이 들어있는 키워드들을 더 많이 검색하고 더 많이 클릭한다는 것을 알 수 있었다.

상권 분석

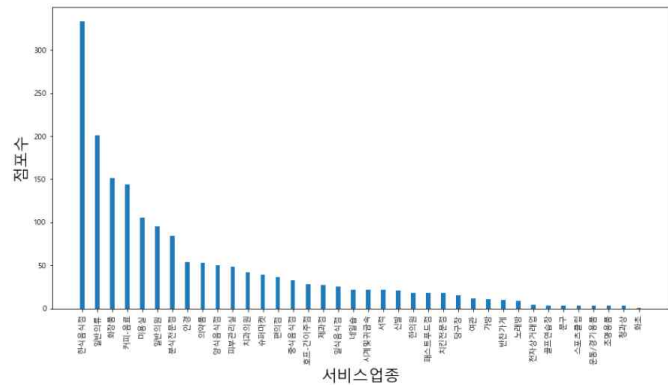
3.1) 업종 분석



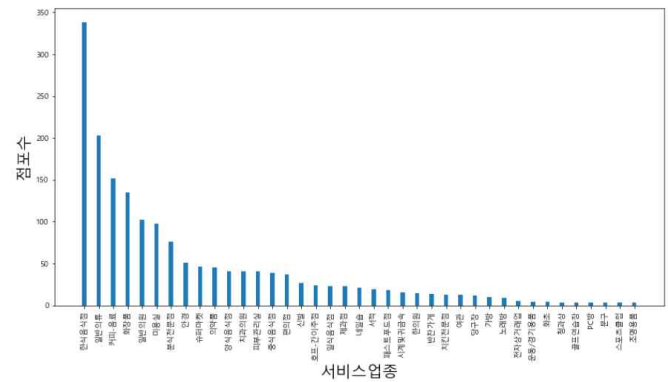
사진[3-1-1]. 21년 1분기 업종 별 점포수



사진[3-1-2]. 21년 2분기 업종 별 점포수



사진[3-1-3]. 21년 3분기 업종 별 점포수



사진[3-1-4]. 21년 4분기 업종 별 점포수

1분기 순위	1분기	2분기	3분기	4분기
한식음식점	1(-)	1(-)	1(-)	1(-)
일반의류	2(-)	2(-)	2(-)	2(-)
화장품	3	3(-)	3(-)	4(1▼)
커피·음료	4	4(-)	4(-)	3(1▲)
미용실	5	5(-)	5(-)	6(1▼)
일반의원	6	6(-)	6(-)	5(1▲)
분식점	7	7(-)	7(-)	7(-)
안경	8	8(-)	8(-)	8(-)
양식음식점	9	9(-)	10(1▼)	11(1▼)
의약품	10	10(-)	9(1▲)	10(1▼)

표[3-1-1]. 21년 업종 별 점포수 순위 변화

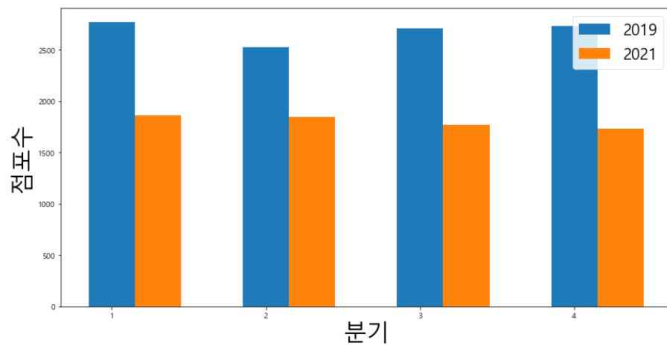
1분기 순위	1분기	2분기	3분기	4분기
한식음식점	350	346(4▼)	333(13▼)	338(5▲)
일반의류	213	214(1▲)	201(13▼)	203(2▲)
화장품	187	161(26▼)	151(10▼)	135(16▼)
커피·음료	123	133(10▲)	144(11▲)	152(1▲)
미용실	103	104(1▲)	105(1▲)	98(7▼)
일반의원	102	98(4▼)	95(3▼)	102(7▲)
분식점	87	87(-)	84(3▼)	76(8▼)
안경	64	66(2▲)	54(8▼)	51(3▼)
의약품	54	54(-)	50(4▼)	45(5▼)

표[3-1-2]. 21년 업종 별 점포수 변화

명동 상권의 분석을 위해 명동상권의 업종 그리고 매출을 분석하고자 한다. 자료는 서울시 열린 데이터 광장에서 제공하는 공공데이터를 가지고 본 목차를 진행하는 것으로 한다.

먼저 최근 데이터인 21년 데이터를 가지고 명동 상권에 위치한 점포의 점포수를 분기별 비교를 하였다. 분기별 특성을 살펴보면, 가장 많은 점포를 보유하였던 한식음식점의 경우 21년동안 그 자리를 굳건히 지켰다고 볼 수 있다. 점포수를 기록한 일반의류 업종 역시 소폭 수치가 줄어들었지만 순위를 유지하였다. 1/4분기~3/4분기 3위를 기록하였던 화장품 업종의 경우 4/4분기에 들어, 점포수가 16개 줄어들며 순위가 한 단계 하락하였다. 반면, 1/4분기~3/4분기 4위를 기록하였던 커피음료 업종은 꾸준히 점포수가 상승하며 4/4분기에 3위로 순위가 한 단계 상승하였다. 점포수를 상세하게 살펴보면 2/4분기에는 화장품 업종의 점포가 26개 줄어든 것을 확인할 수 있고 3/4분기에는 커피·음료 업종과 미용실 업종 점포수가 상승한 것을 제외하면 전체적으로 점포수가 하락한 분기라고 할 수 있다. 4/4분기는 업종별 증감의 굴곡이 심한 분기라고 할 수 있으며 이 시기에도 화장품 업종은 점포가 16개 줄어들며 순위가 하락하였다.

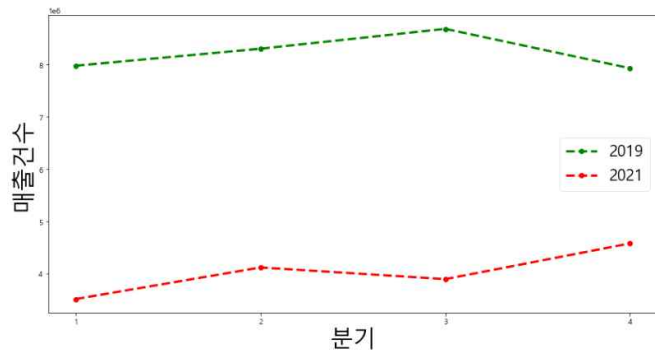
점포수의 분석을 통해 알 수 있었던 것은 화장품 및 한식 음식점 업종의 점포수 감소가 눈에 띄는 부분이 있기는 하나, 순위와 전체 점포수의 추이에 있어서는 정적인 상권이라 할 수 있을 것 같다.



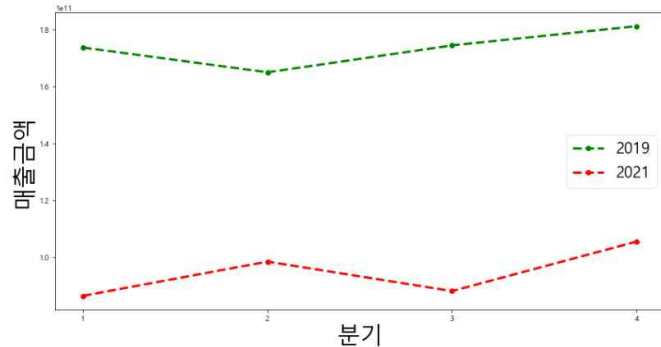
사진[3-1-5]. 2019년과 2021년 점포수 비교

업종 분석 단계의 마지막으로, 다음 그래프는 2019년과 2021년의 분기별 점포수 합계를 시각화한 그래프이다. 다음 자료를 통해 확인 가능한 것은 2019년 그래프는 분기 별 변화 추이에서 큰 연관성을 찾을 수 없었던 반면, 2021년에는 소폭이지만 조금씩 점포수가 감소하는 추세를 보이고 있다는 것이다. 해당 사실이 코로나19 혹은 다른 원인에서 초래된 것인지는 다른 자료를 통해 검증이 필요한 부분이지만, 2019년과 2021년의 점포수 차이가 명확하다는 것은 확실히 확인 할 수 있었다.

3.2) 매출 분석



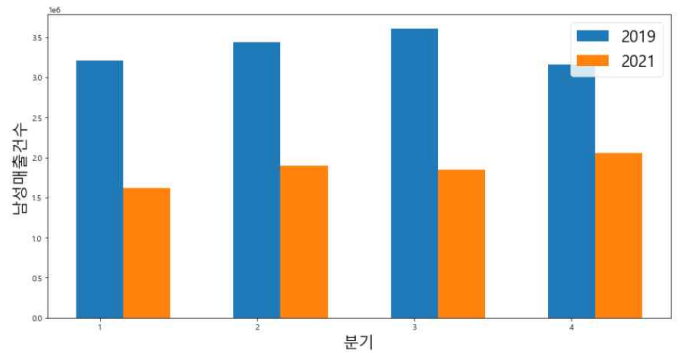
사진[3-2-1]. 2019년과 2021년 분기별 매출 건수 비교



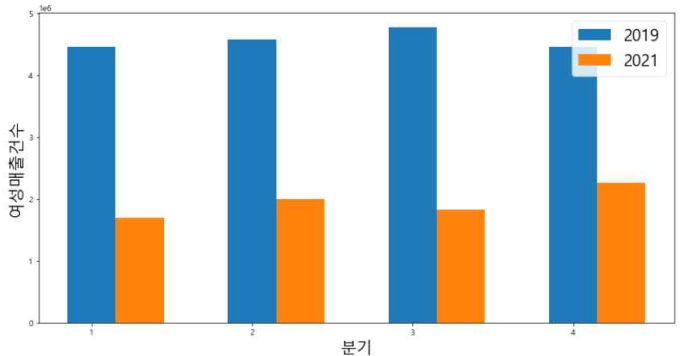
사진[3-2-2]. 2019년과 2021년 분기별 매출 금액 비교

명동 상권의 분석을 위해 두 번째로는, 매출의 분석을 진행한다. 데이터의 비교분석을 위해 2019년과 2021년 두 가지 데이터를 이용하였으며 매출의 건수와 금액 합계를 비교 대상으로 사용하여 그래프를 출력하였다.

매출 건수 그래프를 살펴보면 2019년은 1/4분기부터 3/4분기까지 상승세를 그리다 4/4분기 들어 매출이 급감한 그래프를 나타내고 있다. 반면 2021년 그래프는 4/4분기에 가까워질수록 전반적으로 매출이 상승하고 있다. 앞선 점포수와 마찬가지로 매출 건수에 있어서 2019년과 2021년은 집계된 횟수가 큰 차이를 보인다. 매출 금액 그래프를 살펴보면 2019년은 2/4분기에 매출이 급감한 후, 4/4분기에 가까워질수록 매출이 상승하는 그래프이다. 2021년은 3/4분기에 매출이 급감한 것과 4/4분기에 매출이 급등한 것을 확인할 수 있으며 건수와 마찬가지로 연도 별 금액의 차이가 상당한 것을 확인할 수 있다. 두 그래프를 통해 2019년에는 매출 건수와 매출 금액 사이에 큰 연관성을 찾을 수 없었으며 2021년의 경우, 4/4분기에 매출 건수와 매출 금액 양쪽 모두 크게 상승한 것을 확인할 수 있다.

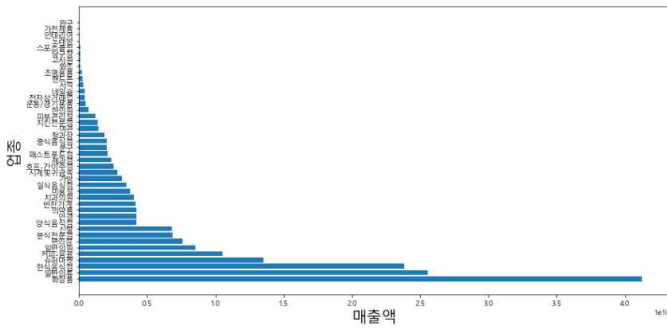


그림[3-2-3]. 2019년과 2021년 분기별 남성 매출 건수

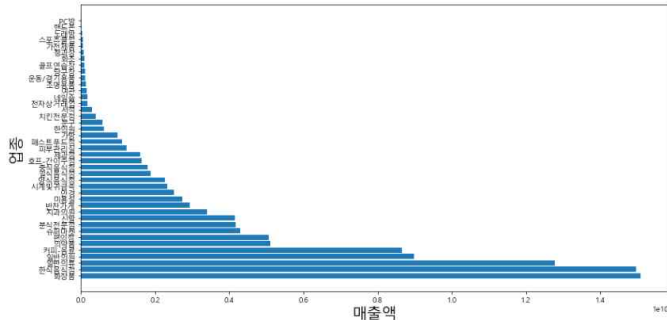


표[3-2-4]. 2019년과 2021년 분기별 여성 매출 건수

다음은 남성과 여성의 분기별 매출건수 집계 그래프를 2019년과 2021년 두 개의 데이터를 이용해 비교 분석을 진행한다. 양쪽 그래프를 보자면 2019년은 남녀 모두 1/4분기~3/4분기에 이르기까지 균일하게 상승하다가 4/4분기에 들어 매출건수가 급감하는 모습이다. 반면, 2021년의 경우 3/4분기에 하락한 것을 제외하면 매출의 건수가 상승한 연도라고 분석 할 수 있다. 이는 앞서 2019년과 2021년의 매출 건수의 비교를 진행하였던 것과 같은 그래프를 그리고 있다는 것을 알 수 있으며 남녀 특성성별에서의 매출 급감이 이루어지지 않고 앞선 데이터와의 일관성이 있음을 보여준다.



사진[3-2-5]. 2019년 업종별 매출 금액



사진[3-2-6]. 2021년 업종별 매출 금액

위 그래프는 2019년과 2021년의 업종별 매출 금액을 비교한 그래프이다. 2019년에 비해 2021년은 매출 금액에 있어서 2019년의 경우 화장품, 일반의류, 한식음식점 3개의 업종이 다른 업종들과 비교해 큰 매출액의 차이를 보이며 상권을 선도해 나갔다고 분석 할 수 있다. 2021년의 경우는 일반의류 업종이 2위에서 3위로 순위가 하락하기는 했지만 1위를 지킨 화장품 업종과 2위로 순위가 상승한 한식음식점은 상위 세 손가락 안을 유지한 것을 확인하였다. 그 외에 일반 의원, 커피음료, 의약품이 상위 업종과의 격차를 줄인것을 확인할 수 있었고 특히, 2019년 12위를 기록한의약품 업종이 2021년에 6위로 순위가 상승한 것을 볼 수 있었다.

1분기 순위	1분기	2분기	3분기	4분기
화장품	1	1(-)	1(-)	2(1▼)
한식음식점	2	2(-)	2(-)	1(1▲)
일반의류	3	3(-)	3(-)	3(-)
일반의원	4	4(-)	4(-)	5(1▼)
의약품	5	6(1▼)	6(-)	7(1▼)
커피음료	6	5(1▲)	5(-)	4(1▲)
편의점	7	7(-)	7(-)	6(1▲)
슈퍼마켓	8	8(-)	9(1▼)	10(1▼)
분식전문점	9	9(-)	8(1▼)	8(-)

표[3-2-1]. 2021년 분기당 매출 금액 변화 추이

명동 상권의 최근 데이터 분석을 위해 2021년 자료를 기준으로 분기별 매출액 순위 변화의 추이를 집계하였다. 2019년과 2021년 그래프의 비교를 통해서도 알 수 있었지만 상위 3개의 업종은 매출액 순위가 분기 당 차이가 크지 않은 것을 알 수 있다. 2021년 1/4분기~3/4분기와 더불어 2019년에도 매출액 1위를

달성하였던 화장품 업종이 21년 4/4분기에 들어 순위가 한 단계 하락한 것과 한식음식점이 1위로 상승한 것을 확인 할 수 있다. 그 외 업종들도 순위 변동은 최대 한 계단 씩 움직이는 것으로 보아 앞으로도 큰 변동은 없을 것으로 예상되고, 커피음료 업종이 꾸준히 순위가 올라가는 것 역시 확인되었다.

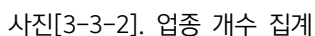
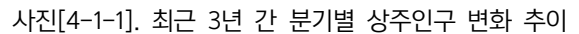
3.3) 상권주도 업종

명동 상권은 조사한 데이터에 따르면 요식업을 하는 업종이 2019년 점포수 집계 데이터에는 상위 10위 안에 4개, 2021년 점포수 집계 데이터에는 상위 10위 안에 3개 그리고 2021년 매출액 집계 데이터에는 상위 10위 안에 3개를 차지할 만큼 그 수와 매출액에 있어서 많은 자리를 차지하고 있고 상권을 주도하는 업종이라고 분석할 수 있다. 따라서 해당 업종을 특정해 분석하기로 한다. 분석에 사용될 데이터는 위와 같은 데이터가 아닌 2022년을 기준으로 '서울시 중구 일반음식점 인허가 정보' 데이터를 통해 상권의 현황 분석을 목표로 진행한다.

조사한 데이터에 따르면 서울시 중구 명동에 위치한 음식점포 가운데 '지번주소'와 '사업장명'이 표기되어 있는 점포는 총 253개 점포인 것으로 확인되었다. 이들 음식점포의 평균 소재지 면적은 약 157미터제곱이고 평균 시설총규모 면적은 약 159미터제곱으로 소재지 면적과 시설총규모 면적의 차이가 크게 나지 않는 것으로 보아 건물밀집도가 높은 지역인 만큼 주차장 및 기타 시설을 갖춘 점포의 비율이 높지 않은 것으로 추측이 가능하다. 가장 큰 면적을 자랑하는 점포는 877미터제곱의 크기로 운영하는 명동2가에 위치한 한식전문점인 '파티파티'이고, 가장 작은 면적을 가진 점포는 명동1가에 위치한 '대학로 수제모찌' 점포이다. 또한 명동을 거점으로 다수의 점포를 운영하고 있는 점포는 '아미소', '명동교자', '유가네 닭갈비' 등 점포로 어느 점포도 3개 이상의 점포를 명동 상권 내에 운영하고 있지는 않은 것으로 파악되었다.



4.1) 상주인구 분석



– 12 –

남성 상주인구 수

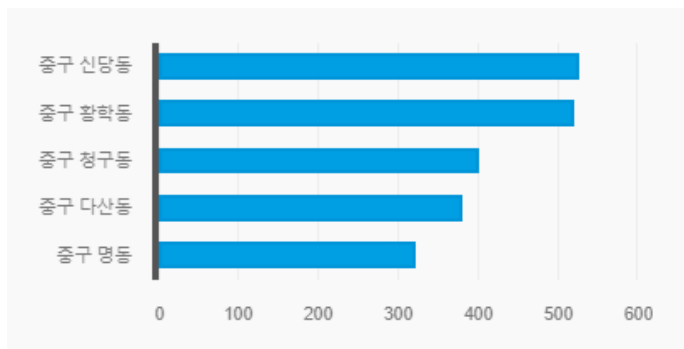


사진[4-1-2]. 2021년 성별별 상주인구 평균

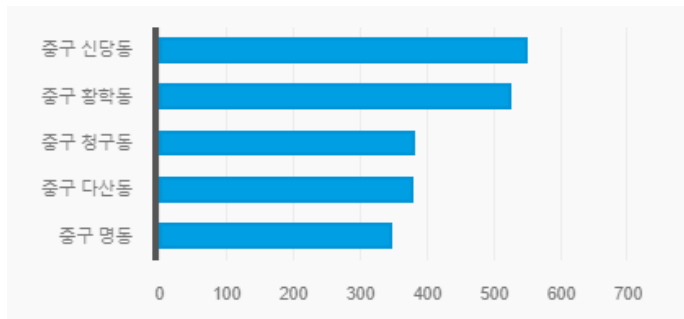
위 그래프도 같은 자료를 통해 추출한 그래프로 2021년 성별별 상주인구의 평균값의 비율을 보여주고 있다. 남성은 상주인구의 약 64%, 여성은 상주인구의 약 36%를 차지하고 있는데, 해당 비율은 2019년과 2020년, 앞선 두 해와 비교해도 그 비율이 크게 달라지지 않았다.

상주인구의 분석을 통해 해당 상권에 상주하는 인구의 숫자와 최근 3년간의 인구 변화 추이 그리고 남성과 여성의 상주인구 비율을 알 수 있었다.

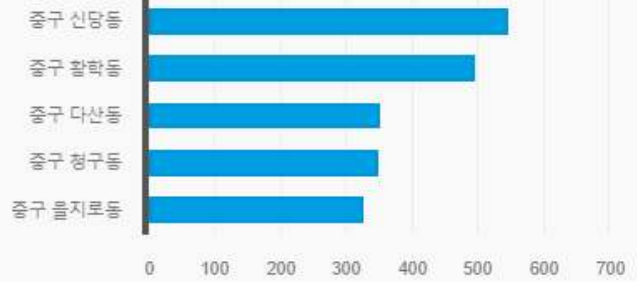
4.2) 유동인구 분석



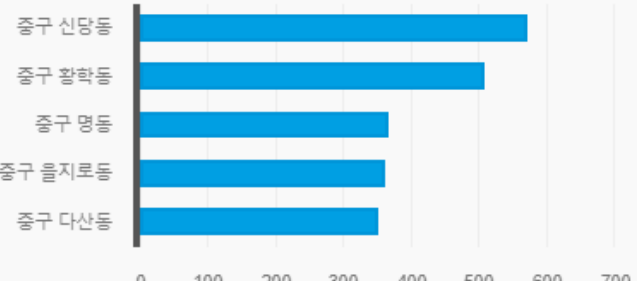
사진[4-2-1]. 2021년 1분기 서울시 중구 유동인구



사진[4-2-2]. 2021년 2분기 서울시 중구 유동인구



사진[4-2-3]. 2021년 3분기 서울시 중구 유동인구



사진[4-2-4]. 2021년 4분기 서울시 중구 유동인구

21년도 서울시 중구에 위치한 행정구역 시군구 단위의 유동인구 집계 데이터를 가져왔다. 명동은 평균적으로 유동인구 350만 명 전후가 집계되었다. 2019년 4/4분기까지 500만 명의 유동인구는 2020년 1/4분기에 들어서 2020년 1/4분기에 들어서는 400만 명 밑으로 떨어지게 되었으며 2018년 이후 처음으로 중구에 위치한 시군구 가운데 5위를 기록하게 되었다. 이는 해당 순위는 2021년 2/4분기까지 유지되어지다, 3/4분에는 5위권 밖으로 밀려났으며 고무적인 부분은 2021년 4/4분기에 들어서 3위까지 순위를 회복했다는 부분이다. 이전의 분석에서 매출과 점포수가 2021년 3/4분에 급감하고 4/4분기에 다시 급증하였다는 분석과 맥락을 같이한다고 볼 수 있으며 앞선 데이터는 2021년 상권이 꾸준히 회복하고 있다는 지표를 보여주고 있으므로 유동인구 역시 기존의 인구를 회복할 것으로 기대가 되며 이는 다시 상권의 매출 증대를 불러일으키는 선순환이 될 것으로 생각된다.

상권 요약

5.1) 상권 분석 정리

○ 지리적 위치

	지역지구	용도지구	용도구역
구분	중심상업지역	방화지구	제1종지구단위계획구역
상권	다양한 업종이 있는 복합상권		

○ 상권 특징

- 유동인구가 많고 여성 유동인구 비율이 더 높은 지역임
- 우리나라 최고 공시지가인 1억9600만원

○ 키워드

□ 워드 클라우드 분석

- ✓ 샵브샵브 전문점인 '정성본', 서울 미쉐린 가이드 2022에 등재된 칼국수 전문점인 '교자' 그리고 현대그룹에서 진행하는 스타트업 프로젝트인 'H 온드림'에 관련된 키워드가 많이 노출되는 것을 확인함

□ 네이버 검색 분석

- ✓ '명동맛집', '명동놀거리', '명동가볼만한곳' 등의 직관적이고 직설적인 표현을 중심으로 많은 검색이 이루어지며 PC와 모바일 양쪽에서의 검색 모두 비슷한 검색형태를 보이는 것을 확인함

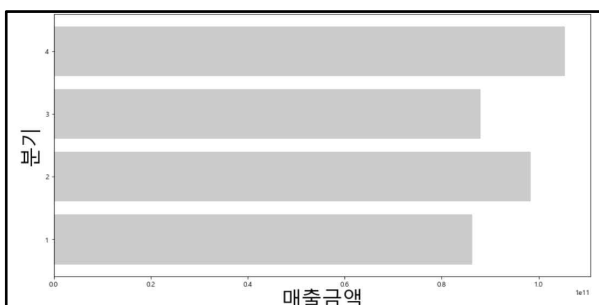
○ 업종분석

점포수	1분기	2분기	3분기	4분기
한식음식점	350	346(4▼)	333(13▼)	338(5▲)
일반의류	213	214(1▲)	201(13▼)	203(2▲)
화장품	187	161(26▼)	151(10▼)	135(16▼)
커피·음료	123	133(10▲)	144(11▲)	152(1▲)
미용실	103	104(1▲)	105(1▲)	98(7▼)

- 2021년에는 화장품, 한식 음식점 점포수가 감소함
- 2019년과 2021년의 점포수는 큰 차이를 명확하게 보임
- 2021년에는 소폭이나마 점포수가 줄어드는 추세임
- 2019년은 화장품 업종, 일반 의류 업종, 한식 업종이 상권을 선도한 것으로 확인함
- 2019년과 비교해 2021년에는 의약품 업종이 6계단 상승한 12위를 기록하면서 일반의원의 꾸준한 매출과 더불어 코로나19의 영향을 간접적으로 보여주고 있음

○ 매출분석

[2021년]

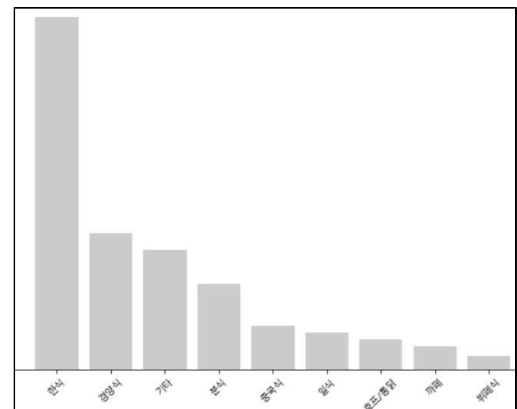


단위 : 원	매출
1분기	86,293,142,126
2분기	98,344,463,805
3분기	87,966,305,362
4분기	105,383,632,978

- 2021년의 매출건수는 3/4분기에 급락하였으며 이는 다시 4/4분기에 급등하였음
- 앞선 분기별 급락과 급등의 추이는 남성과 여성의 성별을 구분한 데이터에도 비슷한 추이를 보이는 것으로 보아 명동 상권에서 특정성별만의 소비 하락은 없었던 것으로 분석함
- 2021년은 앞선 3개의 업종이 선도하던 명동 상권에 커피음료 업종과 일반의원 업종이 가세함
- 2021년 가장 큰 하락세를 보인 업종은 화장품 업종인 것으로 분석함
- 2021년 커피음료 업종은 꾸준한 성장을 기록함
- 화장품 업종과 커피음료 업종 등 몇몇 업종에서의 변화가 있으나 상권 전체적으로 바라보면, 코로나19로 상권 전체가 침체된 것을 제외하고 특정 업종이 크게 급락하거나 순위 변동이 없는 동적인 상권이라 분석함

○ 상권주도 업종

- 명동 상권은 크게 요식업 업종이 주도해 나가는 상권이라 분석함, 요식업종은 어느 연도를 기준으로 삼더라도 매출액이나 점포수 지표가 전체에서 차지하는 비중이 높고 특히 한식 전문점의 비율이 상당히 높음

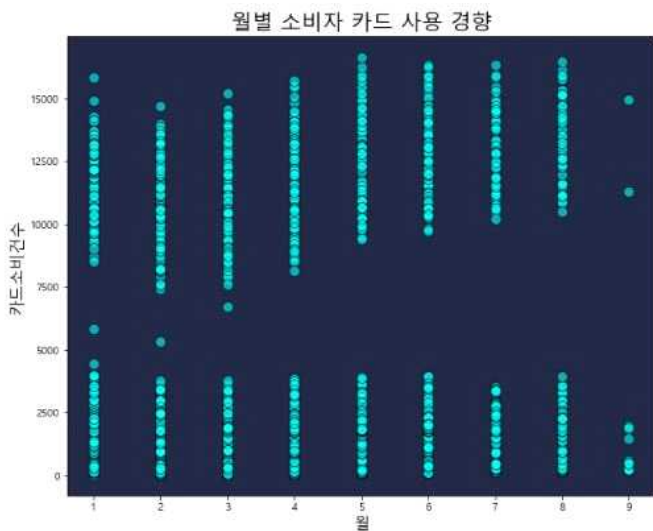


- 요식업 중에서도 한식 전문점의 비율이 높음
- 평균 소재지 면적은 157미터 제곱이고 평균 시설 총규모 면적은 159미터 제곱으로 밀집도가 상당한 지역인 만큼 주차 및 기타 시설이 구비되어 있지 않은 점포가 많다고 분석함
- 명동 상권 내 같은 상호명으로 3개 이상의 점포를 운영하는 점포는 없음
- 명동 상권 내 음식점은 명동2가의 명동4길, 명동8길 그리고 명동1가의 명동7길에 집중분포 되어있음

5.2) 가설 확인

처음에 프로젝트를 시작하면서 “매출액의 상승에 코로나19가 직간접적인 영향을 끼쳤을 것이다.”라는 가설을 두고 상권의 조사 결과에 해당 가설에 대한 내용을 다루겠다고 선언하였고 지금 결과를 말하고자 한다. 일단 기본적으로 넘겨짚고 가야하는 사실은 2021년 3/4분기와 4/4분기는 확진자의 수가 급등하는 사건이 있었고 거리두기의 단계별 해제가 이루어지기 시작한 시점이며 정부의 단계적 일상회복 이행계획 즉, 위드 코로나 프로젝트가 시행되면서 방역패스가 활성화되기 시작한 시점이라는 것을 알아야한다. 따라서, 쟁점은 코로나 확진자 급증, 거리두기의 단계적 해제와 위드 코로나가 시작된 시점인 3/4분기 전후의 데이터가 코로나19와 밀접한 관련성을 보여주고 있는지를 검토해야한다.

첫 번째로, 본문 사진[3-2-1]과 사진[3-2-2]의 2021년 분기별 매출액과 매출건수의 증감 그래프를 살펴보면 1/4, 2/4분기 상승하던 매출과 점포수는 3/4분기에 하락하게 된다. 같은 데이터의 2019년 그래프를 살펴보면 3/4분기로 가면서 매출과 매출건수 양쪽 모두 상승곡선을 그리고 있어, 상권의 특성상 혹은 계절의 특성상 3/4분기가 실적을 내지 못하는 분기라는 의문점은 해소될 수 있다.



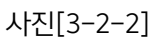
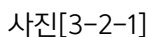
위의 사진은 2019년과 2020년의 데이터를 토대로 머신러닝을 이용하여 카드 소비자의 월별 카드 소비 동향을 업종별, 건수 단위로 산출한 것이다. 해당 그래프를 통해 확인 가능한 사실은 2019년과 2020년 모두 1/4분기, 2/4분기와 비교해 3/4분기에 많은 소비가 이루어짐을 알 수 있는데 이러한 사실 또한 21년 3/4분기에 일어나 매출 급감이 계절 혹은 분기의 특성이 아니라는 것을 뒷받침하는 근거가 된다. 또, 코로나 확진자 수의 급증이 있었던 시기와 일치하며 같은 시기에 명동의 유동인구 순위 역시 같은 중구 내 행정동 순위권에서 멀어지게 된다.(사진[4-2-3]) 따라서 이 시기에 일어난 상권의 침체는 코로나19 확진자의 수 급증과 어느 정도 관련성이 있다고 말할 수 있다.

두 번째로, 정부의 위드 코로나 계획 이행 4/4분기에 유의미한 영향을 끼쳤는지를 확인한다. 표[3-1-2]를 확인하면 3/4분기에는 상위 10개 업종이 대부분 점포수가 하락하지만 4/4분기 들어서 과반 수 이상의 업종이 상승하는 반전을 보인다. 다음으로 표[3-2-1]을 확인하면 분기별 매출액 순위가 4/4분기 들어 다른 분기에 비해 상대적으로 많은 변화를 보인 것을 확인할 수 있으며 자세히 살펴보면 일반의원과 의약품 등 의료 업종과 더불어 화장품 업종만 하락세를 보이고 요식업 업종 위주로 전반적인 상승이 있었다는 것을 확인할 수 있다. 정부의 코로나19 지침 상 요식업 업종의 큰 침체가 있었던 만큼 이러한 상승효과 뒤에는 단계적 일상회복 이행계획이 있었다고 볼 수 있다. 또한, 2021년 3/4분기에 순위권 밖으로 하락했던 중구 행정동별 유동인구 순위는 4/4분기 들어 다시 3위로 회복했다는 것 역시 확인할 수 있었다.(사진[4-2-3]) 매출액과 매출건수 지표의 상승이 연말이라 가능했다는 의문점은 2019년 데이터의 4/4분기 매출건수에서의 하락세를 보인 것으로 해당 반론은 큰 의미가 없을 것으로 생각된다.

구분	인원
경주 신당동	540
경주 칠학동	490
경주 다산동	350
경주 흥주동	340
경주 을지로동	320

[illegible]

5.3) 가설 확인 참고 자료



5.4) 우수점포 소개

(1) 교자 명동 본점



명동 교자는 요식업과 의류업을 중심으로 하는 명동 상권 중에서도 요식업을 선도하는 점포라고 할 수 있다. 1966년 창업했다고 알려진 명동 교자는 현재 명동 내에서 본점과 분점 2개의 점포를 운영하고 있으며 이태원에는 직영점을 운영하고 있다. 우리나라는 전통적으로 메밀의 재배가 발달하여 메밀 국수 의주의 국수가 발달하였고 냉면 역시 그러한 종류의 국수 중 하나라고 할 수 있다. 따라서 우리나라의 밀가루 국수의 역사는 인접한 국가들과 비교해 길지 않다. 그러나 메밀 중심의 국수 판도에 변화가 생긴 일이 있었는데 바로 한국 전쟁이다. 6.25 이후, 미국과의 교류가 늘어나면서 값 싼 밀가루가 대량으로 국내에 유입되었고 이때부터 밀가루를 기반으로 한 요리가 발달하기 시작하였다. 이후, 명동에 ‘장수장’이라는 상호명으로 칼국수를 판매하기 시작한 점포가 바로 현재 명동 교자의 효시인 장수장이다. 이후, 장수장은 상호명은 명동 칼국수로 상호변경을 하였는데, 명동 교자의 정식 개업일은 해당 날짜인 1970년 1월 20일로 되어있다. 장수장이 상호명이었던 시절에도 인기가 많았으나 상호 변경 이후 점포의 인기는 폭발적으로 늘어났고 해당 상호를 벤치마킹하여 ‘명동 칼국수’ 혹은 ‘000 칼국수’라는 이름으로 개업하는 점포들이 많아졌고 상호명의 난립으로 브랜드 이미지가 안좋아지게 이르렀다. 사람들은 명동 칼국수라고 하면, 값 싼 칼국수를 연상하는 지경에 이르렀고 명동 칼국수는 현재의 ‘명동 교자’로 상호명을 변경하게 되고 이는 현재까지 계속 되고 있다.

현재 역시 닭 육수와 애호박, 매운 양파를 기본 베이스로 하는 닭 칼국수를 주 메뉴로 하고 있으며 추가로 만두, 비빔국수, 콩국수 등의 메뉴가 새롭게 출시되었다. 2018년에는 이태원에 직영점을 개업하면서 명동 외에 지점을 두지 않는다는 원칙을 깨버리게 됐다

데, 이는 많은 사람들에게 호평을 받았다. 명동 교자는 미쉐린 가이드 서울 2022에 나오게 되면서 전통과 인기뿐만 아니라 지금 사람들의 입맛에 맞는 상품을 판매하고 있다는 것을 증명하였다.

(2) 롯데호텔 서울점

롯데호텔 서울의 주소지는 서울특별시 중구 을지로 (을지1가)로 명동에 위치하고 있지 않다. 그렇지만 명동의 도보 거리에 인접하고 있으며 명동을 포함해 명동 인근 지역 가운데서 가장 규모가 큰 호텔인 만큼, 롯데호텔 서울 투숙객의 수는 명동 상권에 직접적인 영향을 끼친다고 할 수 있다. 롯데호텔 서울은 그 지리적인 이점과 규모 덕에 과거 롯데 그룹 본사의 기능을 하였던 장소이기도 하며, 현재도 롯데 백화점 본점과 명품관인 에비뉴엘, 영 프라자가 위치하고 있다. 앞서도 얘기했지만 국내에 위치한 호텔 가운데 가장 큰 규모를 가지고 있는데 옛 국립 중앙 도서관 부지를 롯데 그룹이 인수하면서 1978년 12월 본관을 개관하였고, 서울 하계 올림픽 직전인 1988년 6월에 신관을 개관하였다. 명동 등 관광지와 중심지와의 지리적 위치는 외국인 관광객들과 비즈니스맨들에게 큰 이점이며, 롯데호텔 서울의 건설과 증설은 명동 상권에의 엄청난 이득으로 작용하였다. 명동 상권과는 별개로, 호텔 본관 38층에 한식 레스토랑 ‘무궁화’, 뷔페 ‘라세느’ 그리고 미쉐린 2스타를 받은 ‘피에르 가니에르’ 등의 매장이 위치하고 있으며 8년 연속으로 서울 최고의 비즈니스 호텔로 선정되었다.



6.1) 개선사항 및 소감

프로젝트 보고서를 완성하고 난 이후, 다시 보고서를 검토하며 느꼈던 아쉬웠던 부분은 조금 더 다양한 함수와 다양한 기능을 다룰 수 있었더라면 상권을 분석하는데 있어서도 조금 더 구체적이고 직관적인 자료를 구축할 수 있었을 것이라는 생각과 가설을 입증하는데 있어서도 더 논리성과 개연성이 있는 결과물이 나왔을 것이라는 아쉬움이 들었습니다. 코드를 만들면서 어떠한 자료를 특정 조건에 부합시켜 시각화 자료를 만들어낸다면 더 짜임새 있는 보고서가 될 것이라 생각하면서도 원활하게 코딩이 되지 않아 답답했던 시간들도 많았습니다. 그렇지만 포기하지 않고 결과물을 완성해낸 시간동안 많은 것을 배웠다고 생각합니다. 최종 프로젝트뿐만 아닌, 빅데이터 수업을 들으며 데이터 분석의 기초와 파이썬을 이용한 데이터 분석의 여러 기능을 처음 접하였고 배울 수 있었습니다. 또한, 상권 분석이라는 처음 받는 과제들을 수행하며 혼자 그리고 팀원들과 어려움을 헤쳐 나가는 과정도 즐거움이었던 것 같습니다. 제가 보기에다 많이 부족한 결과물이고 만족되지 않는 보고서이지만, 지금 저의 수준에서는 많이 노력한 결과물이라 생각하고 있으며 데이터와 관련된 분야에 흥미가 있는 만큼, 이번 수업과 프로젝트를 발판 삼아 더 높은 코딩 실력과 데이터 분석 능력을 갖추기 위해 노력하겠습니다.

기타

7.1) 코드 설명

1.1) 상권 선정

위의 두 그래프는 '서울시 우리마을가게 상권분석 서비스(상권-추정매출)'의 2019년과 2021년 CSV 형식을 자료를 이용해 출력하였습니다. 상권코드명(상권명)별 그룹핑을 진행하여 분기당 매출 금액을 출력하였고 이를 다시 `mean()` 함수를 통해 평균값 통계 처리를 진행하였습니다. 그리고 이를 다시 분기당 매출 금액을 기준으로 정렬하고 `head(10)`을 이용해 상위 10개 인덱스만 출력한 것을 시각화하였습니다.

1.2) 상권의 지리적 위치

해당 부분은 자료조사 위주로 보고서를 작성하였습니다.

1.3) 상권의 특징

해당 부분은 자료조사 위주로 보고서를 작성하였습니다.

2.1) 워드 클라우드 분석

워드 클라우드의 분석은 네이버 개발자를 이용해 클라이언트 ID와 클라이언트 secret ID를 부여받고 해당 API로 블로그와 뉴스의 검색기록의 제목을 JSON 형태로 추출하였습니다. 이를 다시 워드 클라우드 형태로 시각화해낸 것을 분석하여 보고서를 작성하였습니다. 검색 대상이 되었던 '명동', '명동맛집' 등의 단어들은 모두 같은 코드로 해당 부분만 수정해서 결과를 출력하였습니다.

2.2) 검색 분석

네이버 광고 서비스에 개인 사업자로 등록을 해서, 검색되는 키워드의 양 및 연관 키워드 자료를 엑셀 파일로 받아볼 수 있었습니다. 이를 CSV형식으로 변환하는 과정을 거쳐 pandas에 불러왔고 1차적으로 필요 없는 열을 삭제하는 과정을 거쳤으며 2차적으로 자료의 통계처리를 위해 문자열을 정수형으로 변환하여야만 했습니다. 문자열 가운데 쉽표가 포함된 자료

들의 쉽표를 제거하기 위해 `srt.replace()`함수를 이용하였고 그 후 `astype()`함수를 이용해 문자열을 정수형으로 변환, 통계처리를 진행하였습니다. 통계처리 과정을 통해, 시각화 한 그래프와 데이터프레임을 출력하였습니다.

3.1) 업종 분석

업종 분석은 기본적으로 상권 코드 명을 기준으로 명동인 조건에 해당되는 서비스 업종만을 대상으로 분석을 진행하였습니다. 그래프 출력은 `sort_value`를 이용한 오름차순 정렬과 데이터프레임의 자료를 이용해 출력하였으며 평균값 및 합계에는 기본적인 `mean()`, `sum()`함수 등을 사용하였습니다.

3.2) 매출 분석

매출 분석도 업종 분석과 동일하게 명동에 해당되는 지역만을 선별해 통계처리를 진행하였고 연도 및 분기별 데이터 분류는 그룹핑 처리를 한 이후에 통계처리를 하였습니다.

3.3) 상권주도 업종

음식점포 분석에는 앞선 데이터와 다른 데이터를 사용하여 결측치가 있었고 `dropna()`함수를 이용해 결측치가 있는 컬럼을 삭제한 후 데이터처리를 진행하였습니다. `contains()`함수를 통해 '지번주소'에 명동이 포함되어 있는 데이터를 선별하였고 `max()`, `mean()`함수를 이용해 면적의 최대치와 평균값을 계산하였습니다. 업소의 위치를 지도 위에 표기하는 과정은 `scatter`를 사용하였습니다.

4.1) 상주인구 분석

상주인구의 분석은 간단한 그래프와 원형 그래프의 출력을 통해 진행하였습니다. 자료가 최근 몇 년 간의 자료가 통합되어 있는 데이터였고 연도별 그룹핑을 진행한 후에 다시 분기별 분류를 하였고 통계처리, 시각화를 진행하였습니다.

4.2) 유동인구 분석

해당 부분은 자료조사 위주로 보고서를 작성하였습니다.

5.3) 우수점포 소개

해당 부분은 위주로 보고서를 작성하였습니다.

7.2) 참고 및 활용 자료

(1) 서울 열린 데이터 광장(<https://data.seoul.go.kr/>)

- 서울시 우리마을 상권분석 서비스(상권-추정매출)_2019

- 상권 선정
- 사진[3-1-5]
- 사진[3-2-1]
- 사진[3-2-2]
- 사진[3-2-3]
- 사진[3-2-4]
- 사진[3-2-5]

- 서울시 우리마을 상권분석 서비스(상권-추정매출)_2021

- 상권 선정, 상권 분석
- 사진[3-1-1]
- 사진[3-1-2]
- 사진[3-1-3]
- 사진[3-1-4]
- 사진[3-2-1]
- 사진[3-2-2]
- 사진[3-2-3]
- 사진[3-2-4]
- 사진[3-2-5]
- 사진[3-2-6]
- 표[3-1-1]
- 표[3-1-1]
- 서울시 우리마을 상권분석 서비스(상권-상주인구)
- 상주 인구 분석
- 사진[4-1-1]
- 사진[4-1-2]

- 서울시 중구 일반음식점 인허가 정보

- 상권 분석, 상권주도 업종
- 사진[3-3-1]
- 사진[3-3-2]

(2) 네이버지도(<https://map.naver.com/>)

- 사진[1-2-1], 자료 인용

(3) 이음(<http://www.eum.go.kr/web/am/amMain.jsp>)

- 사진[1-2-2], 자료 인용
- 사진[1-2-3], 자료 인용

(4) 네이버 개발자 - 네이버API(<https://developers.naver.com>)

- 키워드(워드 클라우드) 분석에 사용, 블로그/뉴스
- 사진[2-1-1]

- 사진[2-1-2]

- 사진[2-1-3]

- 사진[2-1-4]

- 사진[2-1-5]

- 사진[2-1-6]

(5) 네이버 광고 서비스(<https://searchad.naver.com>)

- 검색 분석에 사용
- 사진[2-2-1]
- 사진[2-2-2]
- 사진[2-2-3]
- 사진[2-2-4]

(6) 우리마을가게 상권분석 서비스(<https://golmok.seoul.go.kr>)

- 유동인구 조사에 사용, 자료 인용
- 사진[4-2-1]
- 사진[4-2-2]
- 사진[4-2-3]
- 사진[4-2-4]

(7) 실제 사진

- 직접 촬영

(8) 명동교자 (<http://www.mdkj.co.kr/info/info.php>)

- 우수점포 소개

(9) 롯데호텔 (<https://www.lottehotel.com/seoul-hotel/>)

- 우수점포 소개

7.3) 전체코드

전체 코드 가운데 일부를 부분수정 후 처리를 하고 결과를 얻을 수 있는, 중복 되는 부분들이 많은 코드는 해당 문서에 삽입하지 않았습니다.

① -----

```
import pandas as pd
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-추정매출).csv', encoding='cp949')
df2=pd.read_csv('서울시_우리마을가게_상권분석서비스(상권_추정매출)_2019년.csv', encoding='cp949')
df21=df.groupby(['상권_코드_명']).sum()
df2021=df21.sort_values('분기당_매출_금액', ascending=False)
df2021=df2021.reset_index()
df19=df2.groupby(['상권_코드_명']).sum()
df2019=df19.sort_values('분기당_매출_금액', ascending=False)
df2019=df2019.reset_index()
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
plt.figure(figsize=(14,7))
plt.bar(df2019['상권_코드_명'].head(10), df2019['분기당_매출_금액'].head(10))
plt.xlabel('매출', size=20)
plt.ylabel('상권명', size=20)
plt.title('2019년 서울 상권 분기별 매출 합계', size=30)
plt.grid(alpha=0.5)
plt.xticks(rotation=45)
plt.show()
plt.figure(figsize=(14,7))
plt.bar(df2021['상권_코드_명'].head(10), df2021['분기당_매출_금액'].head(10))
plt.xlabel('매출', size=20)
plt.ylabel('상권명', size=20)
plt.title('2021년 서울 상권의 분기별 매출 합계', size=30)
plt.grid(alpha=0.5)
plt.xticks(rotation=45)
plt.show()
df2019[(df2019['상권_코드_명']=='가락시장')]['분기당_매출_금액']
df2021[(df2021['상권_코드_명']=='가락시장')]['분기당_매출_금액']
```

② -----

```
from bs4 import BeautifulSoup
import requests
import pandas as pd
import json
parameter={
    'query':'명동거리',
    'display':'100',
    'start':'1',
    'sort':'date'
}
```

```

}
myheader={
    'X-Naver-Client-Id':'7XEjEKMkrgly40G6nVDv',
    'X-Naver-Client-Secret':'O1iJEa4pz8'
}
data=requests.get(url, params=parameter, headers=myheader).json()
result=[]
start=1
end_point='https://openapi.naver.com/v1/search/news.json?start='
url=end_point+str(start)
data=requests.get(url, params=parameter, headers=myheader).json()
while start<=1000 :
    result=result+data['items']
    start=start+int(parameter['display'])
    #print(start)
    data=requests.get(url, params=parameter, headers=myheader).json()
df=pd.DataFrame(result)
result=""
for i in range(1000) :
    result=result+df['title'][i]
result=' '.join(df['description'])
import matplotlib.pyplot as plt
from wordcloud import WordCloud
result=result.replace('b',"")
wc= WordCloud(font_path='c:/windows/fonts/malgun.ttf',
               background_color='white').generate(result)
plt.figure(figsize=(15,10))
plt.imshow(wc)
plt.axis('off')
plt.show()

```

③ -----

```

import pandas as pd
df=pd.read_csv('명동맛집.csv', encoding='cp949')
df1=df[df['연관키워드'].str.contains('명동')].reset_index()
del df1['index']
del df1['월평균노출 광고수']
del df1['경쟁정도']
del df1['월평균클릭률(모바일)']
del df1['월평균클릭률(PC)']
df1['월간검색수(PC)']=df1['월간검색수(PC)'].str.replace(',','')
df1['월간검색수(PC)']=df1['월간검색수(PC)'].astype(int)
df1['월간검색수(모바일)']=df1['월간검색수(모바일)'].str.replace(',','')
df1['월간검색수(모바일)']=df1['월간검색수(모바일)'].astype(int)
df2=df1.sort_values('월간검색수(모바일)', ascending=False)
df2=df2.reset_index()
del df2['index']
df2
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc

```

```

import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
plt.figure(figsize=(14,7))
plt.bar(df2['연관키워드'].head(10), df2['월간검색수(모바일)'].head(10))
plt.xlabel('검색횟수', size=20)
plt.ylabel('검색 단어', size=20)
plt.title('월간 검색량 통계', size=30)
plt.grid(alpha=0.5)
plt.xticks(rotation=45)
plt.show()

```

④ -----

```

import pandas as pd
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-추정매출).csv', encoding='cp949')
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df1=df[(df['상권_코드_명']=='명동(명동거리)')]
df2=df1[(df1['기준_분기_코드']==1)][['서비스_업종_코드_명','점포수','분기당_매출_금액','분기당_매출_건수','주중_매출_비율']]
df3=df2[['서비스_업종_코드_명','점포수']]
df4=df3.sort_values('점포수', ascending=False)
plt.figure(figsize=(14,7))
plt.bar(df4['서비스_업종_코드_명'], df4['점포수'],width=0.3)
plt.xlabel('서비스업종', size=20)
plt.ylabel('점포수', size=20)
plt.xticks(rotation=90)
plt.show()

```

⑤ -----

```

import pandas as pd
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-추정매출).csv', encoding='cp949')
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df1=df[(df['상권_코드_명']=='명동(명동거리)')]
df2=df1[(df1['기준_분기_코드']==2)][['서비스_업종_코드_명','점포수','분기당_매출_금액','분기당_매출_건수','주중_매출_비율']]
df3=df2[['서비스_업종_코드_명','점포수']]
df4=df3.sort_values('점포수', ascending=False).head(10)

```


⑥

```

import folium
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df=pd.read_csv('서울시 중구 일반음식점 인허가 정보.csv', encoding='cp949')
df1=df.dropna(subset=['지번주소','좌표정보(X)','좌표정보(Y)'])
df2=df1[df1['지번주소'].str.contains('명동')]
df2.info()
plt.figure(figsize=(20,10))
plt.scatter(df2['좌표정보(X)'], df2['좌표정보(Y)'], c='r')
plt.show()
df3=pd.DataFrame(df2['업태구분명'].value_counts())
df3
plt.figure(figsize=(14,7))
plt.bar(df3.index,df3['업태구분명'])
plt.xlabel('업종 개수', size=30)
plt.ylabel('업종명', size=30)
plt.xticks(rotation=45)
plt.show()
df4=df.dropna(subset=['지번주소','소재지면적'])
df5=df4[df4['지번주소'].str.contains('명동')]
df5['소재지면적'].mean()
df5=df.dropna(subset=['지번주소','다중이용업소여부'])
df5['다중이용업소여부'].value_counts()
df6=df5[df5['지번주소'].str.contains('명동')]
df5['다중이용업소여부'].value_counts()
df4=df.dropna(subset=['지번주소','시설총규모'])
df5=df4[df4['지번주소'].str.contains('명동')]
df5['시설총규모'].mean()
df4=df.dropna(subset=['지번주소'])
df5=df4[df4['지번주소'].str.contains('명동')]
df5.info()

```

⑦

```

import pandas as pd
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-추정매출).csv', encoding='cp949')
df2=pd.read_csv('서울시_우리마을가게_상권분석서비스(상권_추정매출)_2019년.csv', encoding='cp949')
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'

```

```

font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df1=df[(df['상권_코드_명']=='명동(명동거리)')]
df3=df2[(df2['상권_코드_명']=='명동(명동거리)')]
df2=df1.groupby('기준_분기_코드').sum().reset_index()
df4=df3.groupby('기준_분기_코드').sum().reset_index()
df2=df2[['기준_분기_코드','분기당_매출_금액']]
df4=df4[['기준_분기_코드','분기당_매출_금액']]
x=[1,2,3,4]
x
[1, 2, 3, 4]
plt.figure(figsize=(14,7))
plt.plot(df4['기준_분기_코드'], df4['분기당_매출_금액'],marker='o', c='green', ls='dashed', lw=3, label='2019')
plt.plot(df2['기준_분기_코드'], df2['분기당_매출_금액'], marker='o', c='red', ls='dashed', lw=3,label='2021')
plt.legend(fontsize=20)
plt.xlabel('분기', size=30)
plt.ylabel('매출금액', size=30)
plt.title("", size=20)
plt.xticks(x, rotation=0)
plt.show()

```

⑧ -----

```

import pandas as pd
import numpy as np
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-추정매출).csv', encoding='cp949')
df2=pd.read_csv('서울시_우리마을가게_상권분석서비스(상권_추정매출)_2019년.csv', encoding='cp949')
df2
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df1=df[(df['상권_코드_명']=='명동(명동거리)')]
df3=df2[(df2['상권_코드_명']=='명동(명동거리)')]
df2=df1.groupby('기준_분기_코드').sum().reset_index()
df4=df3.groupby('기준_분기_코드').sum().reset_index()
df2=df2[['기준_분기_코드','분기당_매출_금액']]
df4=df4[['기준_분기_코드','분기당_매출_금액']]
x=[1,2,3,4]
x
plt.figure(figsize=(14,7))
plt.bar(df4['기준_분기_코드'], df4['분기당_매출_금액'],width=0.3, label='2019')
plt.bar(df2['기준_분기_코드']+0.3, df2['분기당_매출_금액'],width=0.3,label='2021')
plt.legend(fontsize=20)
plt.xlabel('분기', size=20)
plt.ylabel('매출액', size=20)
plt.title("", size=20)

```

```
plt.xticks(x, rotation=0)
plt.show()
```

⑨ -----

```
import pandas as pd
import numpy as np
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-추정매출).csv', encoding='cp949')
df2=pd.read_csv('서울시_우리마을가게_상권분석서비스(상권_추정매출)_2019년.csv', encoding='cp949')
df2
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df1=df[(df['상권_코드_명']=='명동(명동거리)')]
df3=df2[(df2['상권_코드_명']=='명동(명동거리)')]
df1=df1[['기준_분기_코드','서비스_업종_코드_명','분기당_매출_금액']]
df3=df3[['기준_분기_코드','서비스_업종_코드_명','분기당_매출_금액']]
df1=df1.sort_values('분기당_매출_금액', ascending=False)
df3=df3.sort_values('분기당_매출_금액', ascending=False)
plt.figure(figsize=(14,7))
plt.barh(df1['서비스_업종_코드_명'],df1['분기당_매출_금액'])
plt.xlabel('매출액', size=20)
plt.ylabel('업종', size=20)
plt.title("", size=20)
plt.xticks(rotation=0)
plt.show()
plt.figure(figsize=(14,7))
plt.barh(df3['서비스_업종_코드_명'],df3['분기당_매출_금액'])
plt.xlabel('매출액', size=20)
plt.ylabel('업종', size=20)
plt.title("", size=20)
plt.xticks(rotation=0)
plt.show()
```

⑩ -----

```
import pandas as pd
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-추정매출).csv', encoding='cp949')
df2=pd.read_csv('서울시_우리마을가게_상권분석서비스(상권_추정매출)_2019년.csv', encoding='cp949')
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df1=df[(df['상권_코드_명']=='명동(명동거리)')]
df3=df2[(df2['상권_코드_명']=='명동(명동거리)')]
```



```

df2=df1.groupby('기준_분기_코드').mean().reset_index()
df4=df3.groupby('기준_분기_코드').mean().reset_index()
df2=df2[['기준_분기_코드','주중_매출_비율']]
df4=df4[['기준_분기_코드','주중_매출_비율']]
x=[1,2,3,4]
plt.figure(figsize=(14,7))
plt.plot(df4['기준_분기_코드'], df4['주중_매출_비율'],marker='o', label='2019')
plt.plot(df2['기준_분기_코드'], df2['주중_매출_비율'],marker='o', label='2021')
plt.legend(fontsize=20)
plt.xlabel('분기', size=20)
plt.ylabel('주중매출비율', size=20)
plt.grid(alpha=0.8)
plt.xticks(x, rotation=0)
plt.show()

```

⑪ -----

```

import pandas as pd
import numpy
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-상주인구).csv', encoding='cp949')
df
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df1=df[(df['상권 코드 명']=='명동(명동거리)')]
df2=df1[(df1['기준_년_코드']==2019)][['기준_분기_코드','총 상주인구 수','남성 상주인구 수','여성 상주인구 수']]
df3=df1[(df1['기준_년_코드']==2020)][['기준_분기_코드','총 상주인구 수','남성 상주인구 수','여성 상주인구 수']]
df4=df1[(df1['기준_년_코드']==2021)][['기준_분기_코드','총 상주인구 수','남성 상주인구 수','여성 상주인구 수']]
x=[1,2,3,4]
plt.figure(figsize=(14,7))
plt.plot(df2['기준_분기_코드'], df2['총 상주인구 수'],marker='o', label='2019')
plt.plot(df3['기준_분기_코드'], df3['총 상주인구 수'],marker='o', label='2020')
plt.plot(df4['기준_분기_코드'], df4['총 상주인구 수'],marker='o', label='2021')
plt.legend(fontsize=20)
plt.xlabel('분기', size=20)
plt.ylabel('총 상주인구 수', size=20)
plt.grid(alpha=0.8)
plt.xticks(x, rotation=0)
plt.show()
df5=df4[['남성 상주인구 수','여성 상주인구 수']].mean()
df5
plt.pie(df5.values, labels=df5.index, autopct='%0f%%')

```

⑫ -----

```

import folium
import pandas as pd

```

```

import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df=pd.read_csv('서울시 중구 일반음식점 인허가 정보.csv', encoding='cp949')
df1=df.dropna(subset=['지번주소','좌표정보(X)','좌표정보(Y)'])
df2=df1[df1['지번주소'].str.contains('명동')]
df2.info()
plt.figure(figsize=(20,10))
plt.scatter(df2['좌표정보(X)'], df2['좌표정보(Y)'], c='r')
plt.show()
df3=pd.DataFrame(df2['업태구분명'].value_counts())
df3
plt.figure(figsize=(14,7))
plt.bar(df3.index,df3['업태구분명'])
plt.xlabel('업종 개수', size=30)
plt.ylabel('업종명', size=30)
plt.xticks(rotation=45)
plt.show()
df4=df.dropna(subset=['사업장명','지번주소','소재지면적'])
df5=df4[df4['지번주소'].str.contains('명동')]
df5['소재지면적'].mean()
df5['사업장명'].value_counts()
df6=df5.sort_values('소재지면적',ascending=False)
df6[['사업장명','소재지면적']]
df6=df5.sort_values('소재지면적')
df6[['사업장명','소재지면적']]
df4=df.dropna(subset=['지번주소','시설총규모'])
df5=df4[df4['지번주소'].str.contains('명동')]
df5['시설총규모'].mean()
df4=df.dropna(subset=['지번주소'])
df5=df4[df4['지번주소'].str.contains('명동')]
df5.info()

```

⑬ -----

```

import pandas as pd
import numpy
df=pd.read_csv('서울시 우리마을가게 상권분석서비스(상권-상주인구).csv', encoding='cp949')
df
import matplotlib.pyplot as plt
from matplotlib import font_manager, rc
import matplotlib
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df1=df[(df['상권 코드 명']=='명동(명동거리)')]

```

```

df2=df1[(df1['기준_년_코드']==2019)][['기준_분기_코드','총 상주인구 수','남성 상주인구 수','여성 상주인구 수']]
df3=df1[(df1['기준_년_코드']==2020)][['기준_분기_코드','총 상주인구 수','남성 상주인구 수','여성 상주인구 수']]
df4=df1[(df1['기준_년_코드']==2021)][['기준_분기_코드','총 상주인구 수','남성 상주인구 수','여성 상주인구 수']]
x=[1,2,3,4]
plt.figure(figsize=(14,7))
plt.plot(df2['기준_분기_코드'], df2['총 상주인구 수'],marker='o', label='2019')
plt.plot(df3['기준_분기_코드'], df3['총 상주인구 수'],marker='o', label='2020')
plt.plot(df4['기준_분기_코드'], df4['총 상주인구 수'],marker='o', label='2021')
plt.legend(fontsize=20)
plt.xlabel('분기', size=20)
plt.ylabel('총 상주인구 수', size=20)
plt.grid(alpha=0.8)
plt.xticks(x, rotation=0)
plt.show()
df5=df4[['남성 상주인구 수','여성 상주인구 수']].mean()
df5
tf4=df3[['남성 상주인구 수','여성 상주인구 수']].mean()
tf4
tf5=df2[['남성 상주인구 수','여성 상주인구 수']].mean()
tf5
plt.pie(df5.values, labels=df5.index, autopct='%0f%%')

```

⑭ -----

```

import pandas as pd
import matplotlib.pyplot as plt
from sklearn.preprocessing import StandardScaler
from sklearn.cluster import KMeans
import matplotlib
from matplotlib import font_manager, rc
font_location='c:/Windows/fonts/malgun.ttf'
font_name=font_manager.FontProperties(fname=font_location).get_name()
matplotlib.rc('font',family=font_name)
df=pd.read_excel('카드.xlsx')
df
df['이용일자']=df['이용일자'].astype(str)
for i in range(len(df)):
    df.loc[i,'이용일자']=df.loc[i,'이용일자'][4:6]
df['이용일자']=df['이용일자'].astype(int)
df=df.rename(columns={'이용일자':'이용월'})
df['업종대분류']=df['업종대분류'].replace(
    {'가전/가구':0,
     '가정생활/서비스':1,
     '교육/학원':2,
     '미용':3,
     '스포츠/문화/레저':4,
     '여행/교통':5,
     '요식/유흥':6,
     '유통':7,

```

```

'음/식료품':8,
'의료':9,
'자동차':10,
'주유':11,
'패션/잡화':12})
X=df[['이용월', '카드결제건수(천건)']].values
X
kmeans=KMeans(n_clusters=9)
kmeans.fit(X)
kmeans.predict(X)
kmeans.labels_
kmeans.inertia_
result = []
for i in range(1,11):
    kmeans = KMeans(n_clusters=i)
    kmeans.fit(X)
    print(i, kmeans.inertia_)
    result.append(kmeans.inertia_)
df['클러스터번호']=kmeans.predict(X)
plt.figure(figsize=(10,8))
plt.title('월별 소비자 카드 사용 경향', fontsize=20)
ax=plt.gca()
ax.set_facecolor('#212946')
plt.ylabel('카드소비건수', fontsize=15)
plt.xlabel('월', fontsize=15)
plt.scatter(df[df['클러스터번호']==0]['이용월'],
            df[df['클러스터번호']==0]['카드결제건수(천건)'],
            s=100,color='aqua', alpha=.6, edgecolor='black', label = 'Cluster 0')
plt.scatter(df[df['클러스터번호']==1]['이용월'],
            df[df['클러스터번호']==1]['카드결제건수(천건)'],
            s=100, label = 'Cluster 1',color='aqua', alpha=.6, edgecolor='black')
plt.scatter(df[df['클러스터번호']==2]['이용월'],
            df[df['클러스터번호']==2]['카드결제건수(천건)'],
            s=100, label = 'Cluster 2',color='aqua', alpha=.6, edgecolor='black')
plt.scatter(df[df['클러스터번호']==3]['이용월'],
            df[df['클러스터번호']==3]['카드결제건수(천건)'],
            s=100, label = 'Cluster 3',color='aqua', alpha=.6, edgecolor='black')
plt.scatter(df[df['클러스터번호']==4]['이용월'],
            df[df['클러스터번호']==4]['카드결제건수(천건)'],
            s=100, label = 'Cluster 4',color='aqua', alpha=.6, edgecolor='black')
plt.scatter(df[df['클러스터번호']==5]['이용월'],
            df[df['클러스터번호']==5]['카드결제건수(천건)'],
            s=100, label = 'Cluster 5',color='aqua', alpha=.6, edgecolor='black')
plt.scatter(df[df['클러스터번호']==6]['이용월'],
            df[df['클러스터번호']==6]['카드결제건수(천건)'],
            s=100, label = 'Cluster 6',color='aqua', alpha=.6, edgecolor='black')
plt.scatter(df[df['클러스터번호']==7]['이용월'],
            df[df['클러스터번호']==7]['카드결제건수(천건)'],

```



```
s=100, label = 'Cluster 7',color='aqua', alpha=.6, edgecolor='black')
plt.scatter(df[df['클러스터번호']==8]['이용월'],
            df[df['클러스터번호']==8]['카드결제건수(천건)'],
            s=100, label = 'Cluster 8',color='aqua', alpha=.6, edgecolor='black')
plt.scatter(df[df['클러스터번호']==9]['이용월'],
            df[df['클러스터번호']==9]['카드결제건수(천건)'],
            s=100, label = 'Cluster 9',color='aqua', alpha=.6, edgecolor='black')
```