

3세대 골목상권 입주희망자의 매출액 예측을 위한 데이터 셋 구축

Introduction

대한민국은 자영업자의 비율이 타 OECD 대비 유의미하게 높은 편이지만 폐업률마저 높은 편이다. 불안한 고용시장에서 자영업을 선택한 일반 시민들의 결과는 그다지 좋지 못하다는 것을 의미한다. 대부분의 자영업자들이 간과하는 부분은 일단 덮어두고 시작하는 모습이다. 사업을 시작할 때 조금 더 정교한 근거를 가지고 시작하면 폐업률을 낮출 수 있을 것으로 사료된다.

그렇다면 현 시점에서 매출이 높으며 향후 주된 경제활동인구인 20~30대에게 주목해보는 것이 어떨까? 이들을 위해선 기존 알고있던 상권이라는 정의를 조금 바꿔볼 필요가 있다. 과거 전통시장이랑은 확연히 다르며 그렇다고 1세대 격이라 할 수 있는 홍대와 같은 곳은 이미 변질됐다. 사실 이유는 간단하다. 상권에서 버티고 장사할 능력보다 지대를 포함한 투자비용이 비정상적으로 상승했기 때문이다.

이러한 상황에서 그렇다고 장사를 포기하기는 아쉽다. 실제로 데이터가 이를 반증한다. 단순 매출액이 높게 나오는 상가는 강남, 홍대 등이지만 20~30대의 매출액 비중 및 상점당 매출액을 보면 사뭇 결과가 다르다.

무엇이 이들이 기성 상권이 아닌 다른 곳으로 이끌었는지 살펴보자.

근거와 방법은 아래와 같다.





X1, sns현황 - 크롤링

근거 : 특히 3세대 골목상권은 SNS 등 소셜미디어에 의한 자연스러운 전파를 타고 ~ 바이럴마케팅이 아니라 고객, 소비자들이 경험한 바를 SNS를 통해 직접 홍보하며 입소문을 타면서 사람들이 방문하고 이를 기반으로 상권이 형성 - 1.

방법 : sns 언급량 크롤링

X2, 로컬 크리에이터 유무 - 크롤링

근거 : ‘로컬크리에이터’는 골목상권의 주체이자 이해당사자로 민간의 영역에 해당한다.

방법 : 해당 페이지 크롤링으로 크리에이터 지원 사업이 존재하는 경우 1 없으면 0 추가

https://www.k-artup.go.kr//web/contents/webCMRCZN.do?schM=view&id=152658#none

X3, 평당 상가 임대가격 - api, 크롤링

근거 : 비교적 저렴한 임대료를 찾아 새로운 곳을 개척하고 있는 로컬크리에이터의 특성

주소 찾은 후 네이버 부동산 검색 - 상가 면적당 평균 가격 크롤링 - 1.

방법 : 지점을 api 사용으로 주소로 바꾼 뒤 https://land.naver.com/ 에 주소로 검색하면 네이버에서 좌표값으로 자동 변환해 해당 지점에 해당하는 상가 매물 출력됨.

X4, 상권과 지하철 역사와의 거리 - [파일 데이터, api]

근거 : 교통의 중심지, (예시, 종로3가역은 지하철 1ᆞ3ᆞ5호선(종로3가역)이 통과하는 ‘트리플 역세권’) - 1.

방법 : 골목상권의 좌표와 지하철역사의 좌표 간의 거리비교

X5, 버스 정류장과의 거리 - api

근거 : X4와 동일

방법 : X5와 동일

X6, 상권 반경 1km 내 문화시설 거리 - [파일 데이터]

근거 : 문화예술시설을 지원하는 기업이나, 걷기 편한 골목길, 경관이 아름다운 골목상권을 조성하는 정부도 중요한 골목문화의 생산자이다. - 1.

‘나에 대한 탐구는 아마추어 예술활동, 문화예술교육 등 자기표현을 할 수 있는 문화예술 활동을 중심으로 더욱 확대될 것(3.73점)’ - 4.

방법 :

1) 문화예술시설, http://data.seoul.go.kr/dataList/OA-12993/O/1/datasetView.do - 명칭 기반 주소 변환 - 좌표값 변환 후 거리 탐색

2) 서울시 문화공간, http://data.seoul.go.kr/dataList/OA-15487/S/1/datasetView.do - 좌표가 csv에 제공돼있어 거리계산

[혹은 1), 2) 모두 상권을 행정동 단위로 잘라 행정동 내 문화시설 개수로 구하는 방법]

X7, 상가밀도 - [파일 데이터]

근거 : 상가밀도가 낮고 유동인구가 많을수록 매출액이 늘어나는 것으로 확인됐다. -5

방법 : 파일데이터 불러와 전처리

https://data.seoul.go.kr/dataList/OA-15577/S/1/datasetView.do - 상권 내 점포 개수

https://data.seoul.go.kr/dataList/OA-15560/S/1/datasetView.do - 상권 범위

X8, 생활인구 - [파일 데이터]

근거 : 코로나 이후로 전 세계 사람들이 ‘안전’과 ‘거리 두기’를 중시하게 되면서, 생활권의 범위는 어느덧 집을 중심으로 한 ‘동네 상권’에 집중하기 시작했다. -8

방법 : https://data.seoul.go.kr/dataList/OA-15582/S/1/datasetView.do - 배후지역은 상권 주변 200m 이내

X9, 오프라인 클래스 현황 - [크롤링]

근거 : 「두 번째 키워드는 “컨셉진”으로 ‘취향에 맞는콘셉트 세계관 속 콘텐츠로 소통하다’라는 부제목이 있다. MBTI 테스트 등의 성향 테스트를 통해 새로운 관계를 맺거나 기존의 관계를 유지하는 세대, 구성원들과 함께 콘텐츠를 제작하거나 같은 콘텐츠로 놀이를 하는 세대, 타인의 세계관을 존중하는 세대 등으로 설명하고 있다. -9.

방법 : 미정

1. A Case Study on the Local Platform from the Public Design Perspective for the Revitalization of Alley Commercial Area

2. 서울시, 2030세대에 주목…통계로 본 '2030세대'의 삶은?

3. 자영업자의 업종별 매출 영향요인 분석을 위한 머신러닝 기법 활용에 관한 연구

4. 문화예술 트렌드 분석 및 전망 2022-2024

5. 서울시 골목상권 매출액에 영향을 미치는 요인에 관한 연구

6. 골목상권의 매출액과 성장 결정요인

7. MZ세대가 선호하는 형용사별 색채이미지와 주거환경에 대한 기초자료 연구

8. MZ세대의 주거형태와 슬세권

9. 핸드메이드 마켓앱의 사용자 경험연구 -Etsy와 idus를 중심으로-

10. 시계열 군집분석과 로지스틱 회귀분석을 이용한 골목상권 성장요인 연구

