

빅데이터 기반경영 / 사업타당성분석

상권내 업종조합분석

중간발표자료

3조 황희선 남유정 한주엽



목차

01 주제설명

- 프로젝트 목적
- 기존 연구들과의 차별점

02 현재 진행 상황

- 데이터 분석 방법
- 현재 상황 / 문제점

03 향후 계획

04 Q&A



목적

상권 내의 업종별 시너지 분석

코로나와 같은 위기 상황에도
기존의 시너지가 유지되는가에
대하여 검증

코로나 이 후, 또 다른 위기 상황에서
업종별 시너지를 내어 위기 상황을
극복할 수 있도록 도움이 되고자 함

01. 주제설명

기존 연구들과의 차별점

공통점

상권 내의 업종간의
시너지를 분석한다

차별점

시너지 분석을 코로나19라는
위기 상황을 기준 시점으로 나누어
분석한다

02. 현재까지의 진행상황

01	상권내에서 블록별로 나누어 가게들을 장바구니로 구성
02	매출을 기반으로 한 기초적인 긍정/부정 수치 산출
03	업종별 시너지 효과를 제외한 매출에 영향을 줄 수 있는 요소들을 배제
04	코로나 전과 후를 비교하여 코로나에 상관없이 시너지 효과를 보이는 업종끼리 조합
04	결과값을 분석하여 유의미한 다른 결과도 함께 도출

데이터 분석 방법 - 데이터 수집

XML
JSON
소상공인시장진흥공단_상가(상권)정보_API

소상공인 상권정보 상가업소 데이터

5
0
관심

활용신청

오류신고 및
담당자 문의

분류체계	산업·통상·중소기업 - 산업·중소기업일반	제공기관	소상공인시장진흥공단
관리부서명	상권분석실	관리부서 전화번호	042-363-7881
API 유형	REST	데이터포맷	JSON+XML
활용신청	5258	키워드	상가업소,소상공인,상권정보
등록	2015-12-22	수정	2021-11-03
심의유형	개발단계 : 허용 / 운영단계 : 허용		
비용부과유무	무료		
이용허락범위	이용허락범위 제한 없음		
참고문서	소상공인시장진흥공단 상가(상권)정보 활용가이드.zip		

상세기능

주요상권의 영역좌표를 조회하는 기능으로 지정된 상권번호에 해당하는 단일 상권정보를 조회

```

### 서울특별시(시도코드 ctprvnCd = 11) 상권 Data
### 상권명, 위치, 면적

key = "e%2FnG1b5E0R%2Fu0dyYa1Q06U1hB%2FDPkqntCe0gWbyGkNMNy%2BQjM%2BVjy4ub9kfHxt6dEE3Vf%2BUWZ8koFm#BqGcg%3D%3D" # API KEY(일반 인코딩)
url = "http://apis.data.go.kr/B553077/api/open/sdsc2/storeZoneInAdmi?divId=ctprvCd&key={}".format(key) # URL Encoder

content = requests.get(url).content # resultcode 가 00으로 제대로 옴. ※API는 신청하고 약 30분? 지난 후 호출 가능해짐
dict = xmltodict.parse(content)
print(dict['response']['body']['items']) # 실질 데이터

# 파일 저장할 때 jsonString으로 해 줄 수 있다
jsonString = json.dumps(dict['response']['body']['items'], ensure_ascii=False) # 실질 데이터 json으로 변환
jsonObj = json.loads(jsonString)
print(len(jsonObj['item']))

for item in jsonObj['item']:
    print(item)

OrderedDict([('item', [OrderedDict([('trArNo', '9182'), ('mainTrArNm', '종로5가역_3'), ('ctprvCd', '11'), ('ctprvNm', '서울특별시'),
254
{'trArNo': '9182', 'mainTrArNm': '종로5가역_3', 'ctprvCd': '11', 'ctprvNm': '서울특별시', 'signuCd': '11110', 'signuNm': '종로구',
{'trArNo': '9183', 'mainTrArNm': '종로5가역_4', 'ctprvCd': '11', 'ctprvNm': '서울특별시', 'signuCd': '11110', 'signuNm': '종로구',
{'trArNo': '9184', 'mainTrArNm': '종로구청', 'ctprvCd': '11', 'ctprvNm': '서울특별시', 'signuCd': '11110', 'signuNm': '종로구', 'tr
{'trArNo': '9185', 'mainTrArNm': '경모양_1', 'ctprvCd': '11', 'ctprvNm': '서울특별시', 'signuCd': '11110', 'signuNm': '경모양구',
{'trArNo': '9186', 'mainTrArNm': '경모양_2', 'ctprvCd': '11', 'ctprvNm': '서울특별시', 'signuCd': '11110', 'signuNm': '경모양구', 'tr
{'trArNo': '9187', 'mainTrArNm': '경모양_3', 'ctprvCd': '11', 'ctprvNm': '서울특별시', 'signuCd': '11110', 'signuNm': '경모양구', 'tr

```

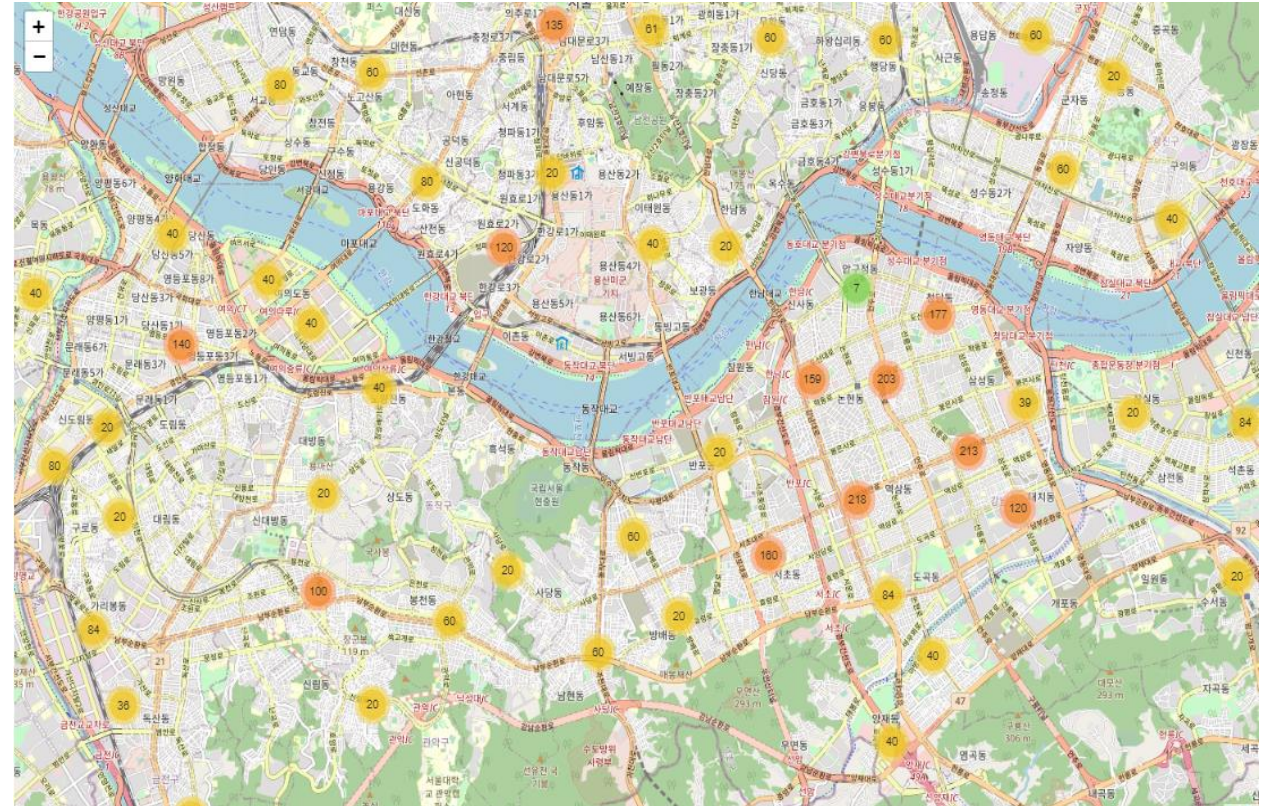

02. 현재까지의 진행상황

데이터 분석 방법 - 데이터 수집

API를 활용한 상권내 상가업소 세부 조회

```
total_stores = pd.DataFrame()
for num in store_df.trarNo:
    request = f'http://apis.data.go.kr/B553077/api/open/sdsc2/storeListInArea?key={num}&serviceKey={key}'
    con = requests.get(request).content
    dict = xmltodict.parse(con)
    # 파일 저장할 때 jsonString으로 해 줄 수 있다
    jsonString = json.dumps(dict['response']['body']['items'], ensure_ascii=False) # 실질 데이터 json으로 변환
    jsonObj = json.loads(jsonString)
    df = pd.DataFrame(jsonObj['item'])
    df['trarNo'] = num
    total_stores = pd.concat([total_stores, df], ignore_index=True)
```

- 지정한 기준에 따라 상권 및 내부 상가 검색 가능



02. 현재까지의 진행상황

매출을 기반으로 하여 기초적인 긍정/부정 수치 산출

- 장바구니 안에 구성되어 있는 업종들의 매출과 업종 평균 매출을 비교
- 장바구니 안의 각 업종의 매출 변화추이(%)가 업종 평균 변화추이(%)에 비해 얼마나 높고 낮은지를 수치로 표현
- 나온 수치를 MinMaxScaler로 표준화하여 0-1사이의 값으로 표현
- 0-1 사이의 값으로 기초적인 긍정/부정 판단
- 빅데이터 캠퍼스에서 제공하는 서울시 업종별 내외국인 신한카드 매출데이터 활용

basket9213 # 9213 상권의 업종 basket

```
[['미/미용/건강', '종식'],  
 ['한식', '종합소매점', '음/식료품소매', '사진/광학/정밀기기소매', '가정/주방/인테리어'],  
 ['음/식료품소매', '종합소매점', '종식'],  
 ['미/미용/건강']]
```

02. 현재까지의 진행상황

데이터 분석 방법 - 상세

업종별 시너지 효과를 제외한 매출에 영향을 줄 수 있는 요소들을 배제



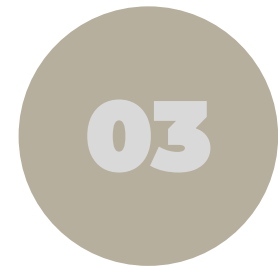
인구

주변 인구의 증가로 인해
자연스럽게 상권의
매출까지 영향을 준 경우



상권 발달 정도

상권 자체의 발달이
급격하여 영향을 준 경우



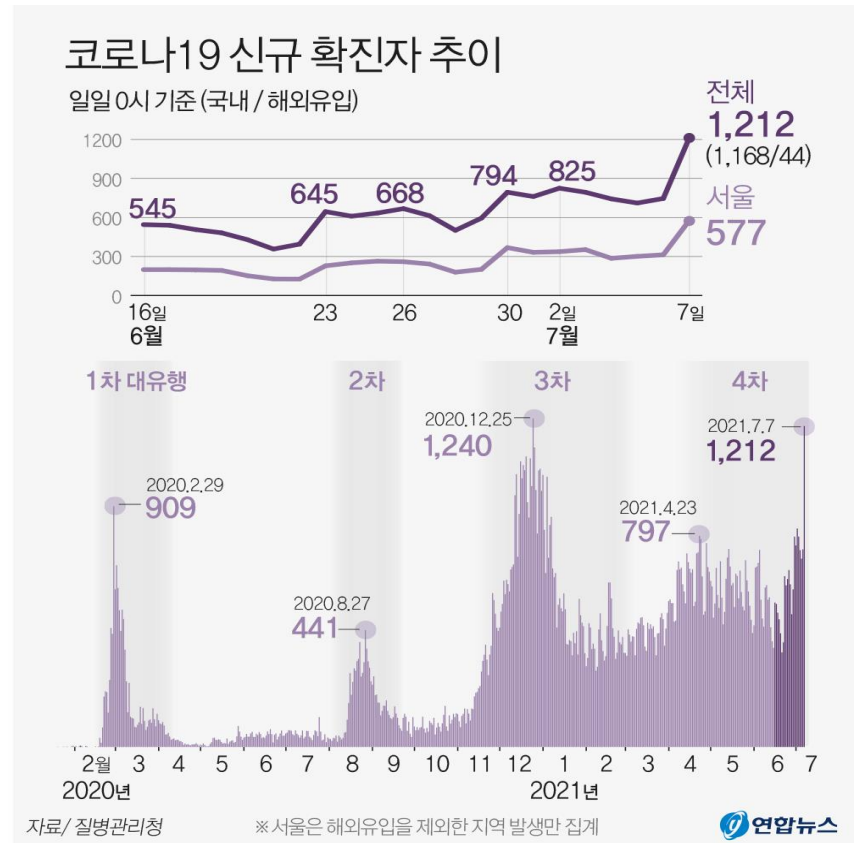
특정 날짜

공휴일, 연휴와 같이 매출의 급격한
증가가 발생하는 날짜

02. 현재까지의 진행상황

데이터 분석 방법 - 기대효과

코로나 전과 후를 비교하여 코로나에 상관없이
시너지 효과를 보이는 업종끼리 조합



- 코로나 발생 전, 1차 대유행, 2차 대유행, 3차 대유행, 4차 대유행을 기준으로 전후를 비교
- 시점에 상관없이 꾸준히 시너지를 발휘하는 업종끼리 조합
- 이 외에도 여러가지 결과값을 분석

02. 현재까지의 진행상황

데이터 분석 방법 - 기대효과

결과값을 분석하여 유의미한 다른 결과도 함께 도출

- 시기별로 분석하였을 때,
여러가지 결과값이 나올 것이라 예상

- 여러가지 결과값을 분석하여
기존에 없던 새로운 결과를 도출해 낼 수 있을 것이라 기대

02. 현재까지의 진행 상황

진행상황/문제점

데이터 수집 및 간단 분석 - 공공데이터API / 공공포털다운로드 / 빅데이터 캠퍼스

공공데이터API

- 우리마을 상권~데이터. api키를 발급, xml형태 -> Pandas Dataframe변환
- 행정구역 단위 상권조회, 상권내 상가업소, 간단한 분석 코드(시각화)

공공포털다운로드

- 상권 인허가 정보(개/폐일)

빅데이터 캠퍼스

- 양재 AI허브 방문
- 블록 단위(현재 상권data)와 카드 매출 데이터(그리드 영역) 확인.

02. 현재까지의 진행 상황

진행상황/문제점

- 블록 단위와 상권의 영역이 다르기 때문에 블록 기준으로 장바구니 영역을 잡았을 때, 상권에 상관없이 장바구니가 구성된다.

데이터가 블록 기준으로 잡혀 있기 때문에, 블록 기준으로 분석하여 해결

- 블록 면적이 워낙 작고 데이터 양이 많다.

우선적으로 마포구를 대상으로 분석을 진행한 후, 점차 늘려가는 방안으로 해결

03. 향후 계획

- 시너지 효과에 대한 함수식의 정리
- 빅데이터 캠퍼스 방문을 통해 데이터를 정리하여 반출
- 반출한 데이터를 분석하고 이를 통해 결과값 도출
- 결과값을 분석하여 새로운 인사이트 추출



Q&A

04

감사합니다.

