

상권-브랜드 이미지 매칭을 통한 창업컨설팅 모형 개선: 커피전문점을 중심으로

1차 보고서(2020.10.13.)

T.O.P. (Toward Optimal Provision)

박시전 권형근 안수빈 이은지 이가영



차 례

- 1 팀 및 주제 소개
- 2 업무 현황 I. 리뷰 크롤링
- 3 업무 현황 II. 상권 구획
- 4 업무 현황 III. 점포별 매출 데이터
- 5 향후 연구 진행 계획



1 팀 및 주제 소개

4주 간의 활동을 거쳐 개별 팀원들의 역할이 일정 수준 확정되었음.

+ 지도교수: 김상용(자문 담당)



박시전

팀장, 경영학과

- 팀 업무 총괄
- 각종 행정서류 담당
- 대외연락 담당 등



권형근

팀원, 컴퓨터학과

- 코드 개발 담당
- 리뷰 크롤링 전담
- 전산 관련 업무 지원



안수빈

팀원, 통계학과

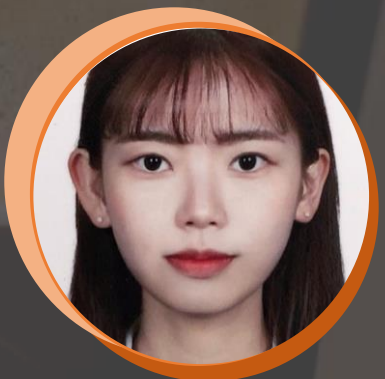
- 데이터 통계 처리 담당
- 상권 구획 업무 총괄
- 유동인구/점포위치



이은지

팀원, 통계학과

- 선행연구 검토,
데이터셋 수집 담당
- 점포별 매출 데이터
업무 지원



이가영

팀원, 통계학과

- 선행연구 검토,
데이터셋 수집 담당
- 점포별 매출 데이터
업무 지원

지난 4주 간 다음과 같이 세 부분으로 나누어 연구 진행을 준비하였음.

주제: 상권-브랜드 이미지 매칭을 통한 창업컨설팅 모형 개선

커피전문점을 중심으로

크롤링 코드 개발

- ✓ 포털 사이트(카카오지도, 네이버지도 등)의 리뷰 창에서 텍스트 형태의 리뷰를 모두 불러오기 위한 코드를 개발하고 수정·보완함
- ✓ 크롤링 결과를 txt 파일로 저장하도록 해 이후 상권 이미지 분석의 기초 데이터로 사용할 것

상권 클러스터링 시도

- ✓ 버스정류장 별 승하차 인원 정보 및 위도 경도 정보를 바탕으로 지역별 유동인구를 파악하려고 함
(상권 클러스터링을 위한 지리적 데이터)
- ✓ 매출액 데이터를 대신해 상가업소 클러스터와 버스정류장 별 유동인구 클러스터의 교집합으로 상권을 새롭게 정의함

데이터셋 탐색

- ✓ 유동인구 데이터, 상가업소 위치 데이터 등은 검색을 통해 쉽게 얻을 수 있었음
- ✓ 그러나 커피전문점 점포별 매출데이터가 얻기 쉽지 않았고, 대부분의 시간이 이 데이터를 조사하는 데 할애됨

2

업무 현황 I. 리뷰 크롤링

(1) 기존의 계획 및 문제점 소개: *스타벅스 양평 DTR점을 예시로*



- ✓ 유명 포털 사이트 및 맛집 정보 사이트에서 **상권의 이미지 분석에 쓰일 데이터** 수집.
- ✓ 상권 내 **각 카페 점포의 리뷰를 크롤링 코드** 개발.
- ✓ **Python**을 통해 페이지의 **HTML 코드**를 **문자열**로 불러온 후, BeautifulSoup 라이브러리를 사용하여 **리뷰 텍스트만 추출**



- ✓ 그러나 동적 웹페이지의 경우 더보기, 다음페이지 버튼 등을 클릭해야만 전체 리뷰 내용이 **HTML 코드**에 반영됨.
- ✓ **Python 내장 함수 만으로는 크롤링 불가!**

2

업무 현황 I. 리뷰 크롤링

(2)-2. 문제점을 어떻게 해결하였는가?

```
while True:
    try:
        button = driver.find_element_by_class_name("_3iTu0")
        button.click()
        time.sleep(1)
        print("%d번째 더보기 클릭"%i)
        i+=1
    except:
        print("더보기 끝")
        break
reviews = driver.find_elements_by_class_name("WoY0w")
print("총 ",len(reviews),"개의 리뷰")
result = []
for review in reviews:
    print(review.text)
    file.write(review.text+'\n')
```

```
for page in range(len(pages)):
    pages = driver.find_elements_by_class_name("link_page")
    try:
        pages[page].click()
    except:
        print("Error")
    driver.implicitly_wait(10)
    time.sleep(3)
    reviews = driver.find_elements_by_class_name("txt_comment")
    for review in reviews:
        print(review.text)
        file.write(review.text+'\n')
```

- ✓ Selenium 패키지의 웹 드라이버를 사용하여 **클릭 이벤트를 처리**.
- ✓ 클릭 시, JavaScript가 동작하여 **동적 웹페이지가 HTML코드에 반영**될 수 있도록 time.sleep(1)을 통해 기다려 줌.
- ✓ Class name을 통해 리뷰 텍스트를 Python에 불러옴.
- ✓ 추후 Python을 이용한 상권의 이미지 분석에서 리뷰 텍스트를 불러오기 용이하도록 **txt파일**에 저장.
- ✓ 각 **포털사이트 별 특징을 반영**하여 코드를 작성하여 **점포별 크롤링은 반복문만으로 해결** 가능하도록 함.

Ex) 네이버(좌측 상단 사진)와 다음 포털 사이트(좌측 하단 사진)에서 '다음 페이지'로 자동으로 넘어가며 지도 내의 지점 리뷰들을 크롤링하는 코드

(3) 현재 상황은 어떠한가?

★아내

직원 불친절. 베이커리 맛있음. 사람 너무 많음. 역대 갔던 스텍 중 최악.

오늘은 날씨까지 참 좋아 강건너 경치 구경하기 좋았어요~

관강 바로 앞에 위치해서 뷰가 끝내준다는 핫한 스타벅스!! 밤이라 풍경은 못봤지만 통유리창으로 된 건 물이 인상적이었고 드라이브스루도 가능해서 ...

통 맛있어요!

통맛있어요. 사람이 생각보다 아직도 많아요. 커피는 동일한 맛. 뷰 포인트는 진짜 일찍 안 오면 없습니다

오픈시간 전에 가도 사람들이 ㅎㅎ했지만 넓고 분위기 좋고 빵도 맛있었음

스타벅스는 어디든 맛있어요

!번째 방문인데, 역시나 아쉬운 부분이 많습니다. 매 장 건축이나 디자인은 새롭지만, 관리 운영은 아쉽 습니다. 1) 매장 청결, 관리가 부족합니다. 사...

정말 낯인대도 사람이 많아요

뷰도 좋고~ 다 좋네요^^ 평일 오후 커피한잔의 여유

커피네요

덕망이 좋네요

방문자가 많아서 정신 없어요 ㅠ 주차나 DT는 안내해주는 분들이 많아 편합니다.

분위기 너무 좋더라고요.

풍경이 멋지고 빵이 맛있고 사람이 많아요

스벅은 진리. 사회적 거리두기 때문인가, 좌석배치도 드문드문. 한가로이 남한강뷰보며 힐

뷰 좋네요~평일 오후라 잘 있다 가요. 사람들 많지 않을 때로 골라서 와야하는 불편함이 있지만, 한적할 거 같은 시간에 또 오고 싶네요. 별적립 관련 문의했는데, 사소한 질문에 여러 직원분들이 친절히 대해주시며 해결해 주었어요. 미소로 응대하는 모습들에 글을 남기게 되네요~^^

어느 스텍을 가도 똑같은 커피맛,매장 청결관리,친절함 스텍은 항상 평타이상은 쳐서 좋음.. 직원분이 추천 해주신 음료 처음 마셔 보는데 만족합니다. 거기서만 파는 '패션푸드 트 캅테일' 음료 시원하고,달달하니 맛있네요. 가족들이랑 잘 마시고 갑니다~

스타벅스에서 볼 수 있는 모든것

조음 사람좀 빠지니까

평일아침에 방문. 일찍부터 사람들이 꽤 있으나, 주말처럼 미친듯이 많지는 않다. 통유리창의 뷰가 좋으며 꽤나 넓다.

좋아요

가주 좋았습니다아아아

사람없고 좋음

그냥 스텍임 아침에 오니까 사람없고 좋네, 빵도 맛남

난웃긴게 뭐만 했다하면 그때들처럼 너도나도 로보트 처럼 하려는 미개한 국민성에 더 소름돋는다 지금 천지로 널린게 스타벅스 매장인데 굳이 한시간주차 한시간 웨이팅 기다리면서까지 커피를 마시려는 이유가 뭐임?거기 커피는 금가루라고 뿌려놨음?오혀려 더 맛없다는데?용진씨도 이런 미개한 국민성을 마케팅으로 삼아 장사하고 있는거지

✓ 예시인 양평DTR점에 대하여 개발한 코드가 각 포털 사이트에서 성공적으로 크롤링함을 확인.

✓ 각 점포별 txt파일을 생성하고, 한 줄에 하나의 리뷰가 들어가도록 저장.

✓ 상기 내용을 통해 본격적인 크롤링 준비 완료됨을 확인.

✓ 내용을 쉽게 Python에 불러올 수 있는 txt 파일로 데이터를 저장하여 추후 상권 이미지 분석에도 매우 용이할 것으로 예상.

✓ But, 각 페이지를 크롤링 하는데에 수초의 대기시간이 발생하므로 대용량 데이터 크롤링의 경우 클라우드 서버를 이용하여 작업을 수행할 예정

Ex) 스타벅스 더양평DTR점 리뷰 크롤링 결과

(좌측 상단: NAVER 지도 내 리뷰, 좌측 하단: DAUM 지도 내 리뷰)

2 업무 현황 I. 리뷰 크롤링

(4) 향후 리뷰 크롤링 작업 완료를 위해 어떻게 할 것인가?

상권 내 점포 URL list-up

- ✓ Google, Naver, Daum 지도 API를 사용하여 자동으로 URL을 list-up 하는 코드를 작성
- ✓ 만약, 위 방법이 안될 경우 직접 점포별 URL을 검색하여 list-up
- ✓ Txt 파일 형식을 이용하여 list-up하고, 추후 Python으로 불러와 반복문을 수행

대용량 리뷰 데이터 크롤링

- ✓ 클라우드 서버를 이용하여 수 천개의 점포에 대한 리뷰 크롤링
- ✓ 결과는 txt파일에 작성하여 데이터 저장
- ✓ 추후 Python을 이용하여 다시 불러와 상권 이미지 분석에 사용

상권의 이미지 분석

- ✓ 수집한 데이터에 대한 분석 진행
- ✓ koNLPy 패키지의 twitter 모듈을 이용하여 형태소 분석
- ✓ 의미 있는 단어들을 도출하여 상권 이미지 분석 및 시각화
- ✓ 결과 저장 format은 이후에 나올 매출 데이터 format에 따라 결정

3 업무 현황 II. 상권 구획

(1) 기존의 계획은 어떠하였는가?

위도&경도 기준 노드 설정

조밀한 직사각형 모양으로 위도와 경도를 분할하여 각 직사각형을 하나의 "노드"로 취급

다차원 클러스터링

위도와 경도, 유동인구, 노드 내 상가 수, 노드 내 상가 총 매출액으로 K-means 기반 5차원 클러스터링 수행

상권 최종 정의

다차원 클러스터링 결과 도출된 여러 노드의 집합을 하나의 "상권"으로 정의

03

02

01

cf. 노드 분할

0.001간격 분할

127.123	127.124	
	37.568	
	37.567	

(2) 발생한 문제점 및 해결방식

문제점

- 개인정보 보호법 및 비용 문제로 인한 매출액 데이터 수집의 어려움
- 임의로 분할 단위를 지정하는 것과 기술적인 문제로 인한 직사각형 모양 노드 생성의 어려움
- 노드 별 정확한 유동인구 및 상점 수 파악의 어려움으로 인한 bias 발생
- 3차원 이상의 다차원 클러스터링은 시각적으로 표현하기 어려움

계획 보완

- 버스정류장별 세부적인 유동인구 파악 가능성 확인 → 직사각형 노드를 나누지 않고 세세한 구역의 유동인구를 파악할 수 있게 됨
- “상권”에 대한 새로운 정의: 유동인구 클러스터&상가업소 클러스터의 “**교집합**” → 유동인구와 상가업소 수가 동시에 많은 지역을 상권으로 정의함으로써 보다 엄밀하게 상권을 규정할 수 있게 됨
- 버스정류장별 좌표 및 유동인구 클러스터링 + 상가업소 좌표 클러스터링 결과 합치기



3

업무 현황 II. 상권 구획

(3) 현재 상황은 어떠한가?

```
df = floating[['X', 'Y', 'f_sum']]
df.head()
```

버스정류장 좌표별(X, Y)
승하차 인원수 합계(f_sum)

	X	Y	f_sum
0	126.987750	37.569765	24294
1	126.996566	37.579183	100495
2	126.998340	37.582671	143669
3	126.987613	37.568579	34281
4	127.001744	37.586243	96051

```
#데이터 표준화_standardScaler
from sklearn.preprocessing import StandardScaler
standardScaler = StandardScaler()
print(standardScaler.fit(df))
df_scaled = standardScaler.transform(df)
```

```
StandardScaler(copy=True, with_mean=True, with_std=True)
```

```
df_scaled
```

```
array([[ 0.03186667,  0.3281433 ,  0.11380605],
       [ 0.13585299,  0.49769442,  2.48739752],
       [ 0.15677858,  0.5604919 ,  3.83222815],
       ...,
       [ 2.22633163,  0.09959713,  0.09686095],
       [ 2.1698711 ,  0.45054865, -0.57324933],
       [ 1.69259514,  0.03205169, -0.50774281]])
```

(1) 유동인구 데이터

각 버스정류장 별 모든 노선에서의 승/하차 인원을 집계하여 전체 유동인구(f_sum) 열을 생성함

=> 클러스터링에서 위도와 경도, 그리고 유동인구 간 단위차에 의한 영향을 없애기 위해 scaling 수행

(2) K-means clustering

```
[13] # 비계층적 군집 분석 k means model
model = KMeans(n_clusters=10, random_state=0, algorithm='auto')
# random_state=0 : seed 역할 (모델을 일정하게 생성 = 랜덤X)
model.fit(df_scaled)
```

```
➡ KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
          n_clusters=10, n_init=10, n_jobs=None, precompute_distances='auto',
          random_state=0, tol=0.0001, verbose=0)
```

```
[14] # 클러스터링(군집) 결과
pred = model.predict(df_scaled)
pred
```

```
➡ array([3, 5, 5, ..., 0, 0, 0], dtype=int32)
```

비계층적 군집 분석 기법인 K-means clustering 수행
+ 버스정류장 위도, 경도, 유동인구의 3차원 클러스터링
+ Hyperparameter: 군집 개수(n_clusters) = 10
+ random_state=0으로 지정하여 모델을 일정하게 생성하도록 제한
+ Scaled data에 대해 modeling

클러스터링 결과, 위와 같이 3번 군집, 5번 군집, 0번 군집 등 다양한 군집에 객체가 할당된 것을 볼 수 있다.

3

업무 현황 II. 상권 구획

(3) 현재 상황은 어떠한가?

[15] # 군집별 중앙값

```
centers = model.cluster_centers_  
centers
```

```
array([[ 1.68866526, -0.49599839, -0.15941029],  
       [-0.75575234, -1.16069633, -0.2853896 ],  
       [-0.83443187, -0.77439955,  1.7041186 ],  
       [ 0.62580489,  0.32570036, -0.2236042 ],  
       [ 0.57832853, -1.12766115, -0.1395578 ],  
       [ 0.46398615,  0.30902592,  2.33713876],  
       [-0.12387005, -0.20499211,  6.65993967],  
       [ 0.74410048,  1.49243021, -0.27679294],  
       [-1.59609965, -0.36106795, -0.303475 ],  
       [-0.58897348,  0.48100911, -0.25741537]])
```

(3) 각 군집별 중앙값 계산

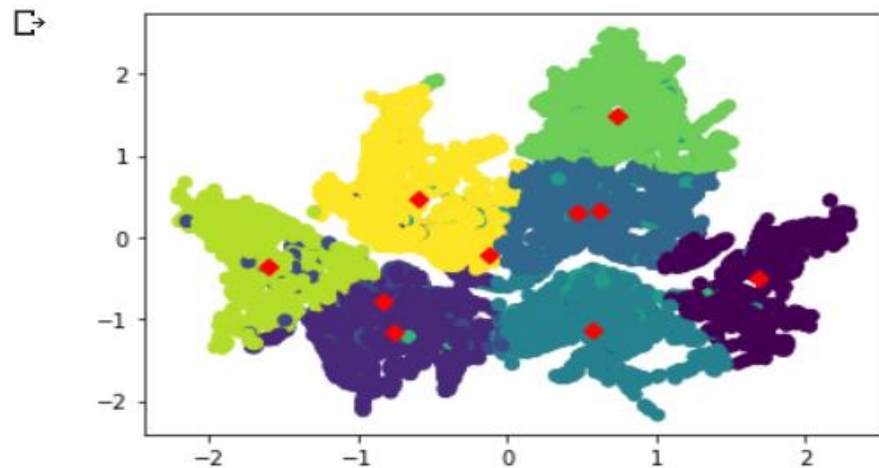
클러스터링 결과 생성된 각 군집에서 **centroid**를 계산한다.

+ 위도의 중앙값, 경도의 중앙값, 유동인구의 중앙값이 도출되므로 **총 3차원의 array**가 생성되었다.

(4) 중앙값과 함께 시각화

[16] # 중앙값과 함께 시각화

```
plt.scatter(x=df_scaled['X'], y=df_scaled['Y'], c=pred)  
plt.scatter(x=centers[:,0], y=centers[:,1], marker='D', c='r')  
plt.show()
```



크게 7개의 유의미한 클러스터가 구분되는 것을 관측할 수 있다.

그 외 3개의 클러스터는 다른 클러스터에 흡수되어 잘 구분되지 않는다. 중앙값을 함께 나타내어보니, 클러스터의 중심과 그 주위에 객체들이 적절하게 할당되었다.

(4) 향후 상권 구획 작업 완료를 위해 어떻게 할 것인가?

유동인구 클러스터링 구체화

- ✓ 3차원 시각화 방법 고안
- ✓ Heuristic한 방법으로 적절한 군집의 수(k) 모색
- ✓ 클러스터링 결과 생성된 각 군집에 대한 위도&경도 범위 파악
- ✓ 최대&최소 위도와 최대&최소 경도로 클러스터 규정 및 csv 생성

=> 추후 상가업소 클러스터링과의 교집합을 도출하는 데에 이용

상가업소 클러스터링 진행

- ✓ 상가업소의 매출 데이터 혹은 대안 데이터 모색
- ✓ 상가업소 좌표(위도, 경도)와 매출 데이터(혹은 대안 데이터)로 클러스터링
- ✓ 생성된 각 군집의 위도&경도 범위 파악
- ✓ 3차원 시각화 및 적절한 군집 수 모색

=> 추후 유동인구 클러스터링과의 교집합을 도출하는 데에 이용

클러스터 교집합 도출

- ✓ 유동인구 클러스터링 결과와 상가업소 클러스터링 결과를 취합해, 중첩되는 지역 좌표의 최대/최소 값 계산
- ✓ 위 결과 도출된 좌표의 범위에 의해 생성된 지역을 “상권”으로 정의
- ✓ 정의된 모든 상권의 리스트를 csv형태로 저장
- ✓ 상권 리스트 csv를 대상으로 지도 위에 2차원(위도, 경도) 시각화

4 업무 현황 III. 점포별 매출 데이터

(1) 기존의 계획은 어떠하였는가?

- ✓ **브랜드 이미지에 대한 사전 연구 확인:** 어떤 기준으로 프랜차이즈 브랜드 표본을 추출할 것인가?
-> 가격이 브랜드 이미지 형성에 유의미한 영향을 미친다는 논문 확인
-> 아메리카노 가격을 기준으로 5개의 브랜드 선정: 스타벅스/백다방/커피빈/투썸플레이스/이디야
- ✓ 서울시 도로명주소별로 도로폭이 나와있는 데이터 확보 후 '도로명' 을 기준으로 구체적인 점포를 선정하고자 함
- ✓ 선정한 몇 군데의 점포들에 대한 프랜차이즈 카페의 지점별 매출 데이터를 얻고자 함



카드사에 전화/메일 연락을 통해 프랜차이즈 점포별 매출 데이터 구매가 가능한지 확인



스타벅스 안암점, 탐앤탐스 안암점 등 안암의 프랜차이즈 카페를 직접 방문해 매출 데이터를 얻을 방안 확인



BC카드 랩실 방문 후 연구 목적/예산에 부합하는 데이터 구매 및 반출 방법 확인



프랜차이즈 본사 PR팀/고객센터에 연락을 취해 매출 데이터 구매 가능 여부 확인

4 업무 현황 III. 점포별 매출 데이터

(2) 현재 상황은 어떠한가?

카드사 연락

100%

BC카드 랩실

70%

프랜차이즈 카페 본사 방문/연락

100%

대안적 지표 모색

40%



카드사는 카페 카테고리에 대한 매출 데이터만 있을 뿐, 프랜차이즈별 매출 데이터도 없음

BC카드 랩실에서는 지역구별로 적합한 데이터 구매 가능 확인
but 재방문 후 반출 절차 거쳐야 함

백다방, 스타벅스, 커피빈 등 대부분 프랜차이즈 본사는
데이터 비공개라는 답을 받음

대안적 지표로 권리금, 임대료 데이터 모색 방안 논의 중

김상용 지도교수님 면담

- 매출데이터는 구하기 어려울테니 대신 '대안적 지표' 활용 추천
- 대안적 지표로는 권리금, 임대료 등을 찾아보면 될 것

유시진 마케팅 교수님 면담

“매출데이터는 구하는 게 거의 불가능할 것. 가게의 performance를 알기 위해서 차라리 몇 군데 가게에 직접 방문하여 점장과 인터뷰를 해보는 게 좋을 것”

유의사항:

민감정보 취급 시
개인정보보호법 등에 유의할 것

4 업무 현황 III. 점포별 매출 데이터

(2) 현재 상황은 어떠한가?

주력 대안

서울시내 BC카드 사용 내역 정보 데이터(커피전문점 한정)

- ✓ 서울시 강남구, 서초구, 송파구, 마포구,
용산구의 카페업종 매출 집계 데이터
- ✓ 가맹점 지역(우편번호)에 따라 분류한
카드 사용 내역 포함
- ✓ 데이터 집계 기간: 1개월 / 일 단위로 제공
- ✓ 주최 측에 데이터 구매를 신청한 상황

보조 데이터

권리금 또는 임대료

- ✓ 영업 중인 카페가 적어도 권리금 및 임대료를 감
당할 수 있을 만큼은 매출을 내고 있다는 가정
- ✓ 온라인으로 조사해보았으나 부족한 점이 많
아, 해당 구의 부동산에 방문하여 추가 자료
수집 시도함
- ✓ 점포별 매출을 지역적 정보와 연관 지어 볼 수
있는 보완적 데이터

4 업무 현황 III. 점포별 매출 데이터

(2) 현재 상황은 어떠한가?

부동산 방문 면담 결과: 대안적 지표로서의 권리금 및 임대료의 문제점

권리금의 문제점

- ✓ 권리금은 매출보다는 상가에 대한 수요, 즉 지역 자체의 부동산 가격에 영향을 받음
- ✓ 카페 폐점 시 권리금을 받고 나가는 구조라서 권리금은 쉽게 공개되지 않는 데이터
- ✓ 점포가 권리금을 받지 않고 나가는 경우도 많음
→ 부동산은 추정치만 제시 가능

임대료의 문제점

- ✓ 동일 지역구 내에서도 상세 주소, 대로변과의 거리 등에 따라 임대료 차이가 큼 (ex. 강남구 대치동의 한티역 주변 / 대치역 주변)
- ✓ 동일 지역 내에서도 준공 시기, 건물 내 점포의 위치(층 수, 가장자리에 있는지 등)에 따라서도 차이가 큼
- ✓ 지역별 임대료를 기준으로 지역의 가치를 추정하기 어려움

4 업무 현황 III. 점포별 매출 데이터

(3) 향후 점포별 매출 데이터 추정 완료를 위해 어떻게 할 것인가?

현 시점, 가장 고민이 많은 지점

주제를 지속할 것인가?

- ✓ 우편번호 단위까지 세분화 되어있는 데이터를 본부 측에 구매요청한 상황
- ✓ 정상적으로 수령하게 되면, 동일한 우편번호 구역 내에서 개별 지점을 특정할 수 있는 방법에 대해 조금 더 고민해볼 예정(11월~)
 - > 권리금/임대료 이외에 또 다른 유효한 보조 데이터의 존재 여부 지속 탐색
 - > 김상용 교수님과의 면담 재요청 예정

주제를 변경할 것인가?

- ✓ 주제 변경의 가능성 또한 이제는 배제하기 어려움
- ✓ 기보유한 데이터셋을 최대한 활용할 수 있는 주제가 무엇이 있을지, 원점에서부터의 재검토 실시
 - > 필요 시 유시진 교수님과의 면담 재요청
- ✓ 최종 주제 변경 여부는 늦어도 11월 20일까지 결정 예정

이 부분에 대한 외부 전문가의 피드백이 필요!

다음 보고 시까지 어느 정도의 업무 진척을 목표로 하는가?

연구 방향성 보완

- ✓ 정확한 점포별 매출액을 알 수 없기 때문에 위치, 유동인구, 노드 내 상가 수, 점포 별 매출액을 기준으로 한 상권 클러스터링이 불가능한 상황

→ 카드 사용 내역으로 얻은 지역별 매출액 데이터는 상권의 수익성을 추정하기 위한 지표로만 사용

→ 상권 클러스터링을 위한 기준 보완 필요

데이터 전처리

- ✓ 분석 대상이 될 커피 프랜차이즈 브랜드 선정
- ✓ 크롤링한 리뷰 데이터에 자연어 처리 적용 및 토큰화
- ✓ 카드 사용 내역 데이터의 우편번호를 위도, 경도로 변환해, 지역별 카페 매출액 정보와 유동인구 데이터가 지리적 정보를 나타내는 열로 묶일 수 있도록 할 것

분석 기법 결정

- ✓ LDA(*Latent Dirichlet Allocation*)와 LSA(*Latent Semantic Analysis*) 중 브랜드 이미지 추출에 더욱 효과적인 기법 결정
- ✓ 브랜드 이미지로 도출된 어휘와 상권 이미지로 도출된 어휘 간의 일치도 등을 이용해 브랜드와 상권 이미지가 매칭되는 정도에 대한 지표 생성

End of the Presentation

T.O.P. (Toward Optimal Provision)

1차 진행 상황 보고 자료

제출일자: 2020.10.13.

발표자: 권형근, 이은지