



망원은 인스타그램이 좋아요

#김지현 #박소연 #최승원 #차영희



강윤철 교수님
최병철 교수님
김승범 교수님



빅데이터기반경영/
사업타당성분석
Project Final Presentation



1

주제 선정 배경 및 문제 제기

- 인스타그램의 영향력 증가 => 망원: '인스타그램 상권'
- 연도 별로 달라지는 인스타그램과의 영향 관계 검증 필요

2

데이터 분석 + 결과 해석(시사점)

- ANOVA, 회귀분석, Stepwise 전진선택법, Interaction Plot 사용
- 데이터 분석 결과 해석 및 시사점 제공

3

분류 분석 모델

망원의 모든 업종을 대상으로 인스타그램과의 영향관계 파악

1

주제 선정 배경 및 문제 제기

- 인스타그램의 영향력 증가 => 망원: '인스타그램 상권'
- 연도 별로 달라지는 인스타그램과의 영향 관계 검증 필요



상권의 잠재력
접근 가능성
성장 가능성
중간 특성
누적적 흡인력
양적성
경쟁 회피성
경제성
...



[토요워치] 井(우물)안 상권도 #(해시태그) 달면 매출 쑥

외진 골목·산꼭대기 가게도 SNS서 입소문
송리단길·망리단길·목리단길 핫플레이스로
권리금 2년새 3,000만원 올라도 매물 없어

출처: 서울경제

핫플레이스가 '인스타'에 찍힌다

이유진 여성신문 기자 | 승인 2017.02.21 10:51 | 수정 2017-03-08 10:40 | 댓글 0

서울 명소 지도 바꾼다

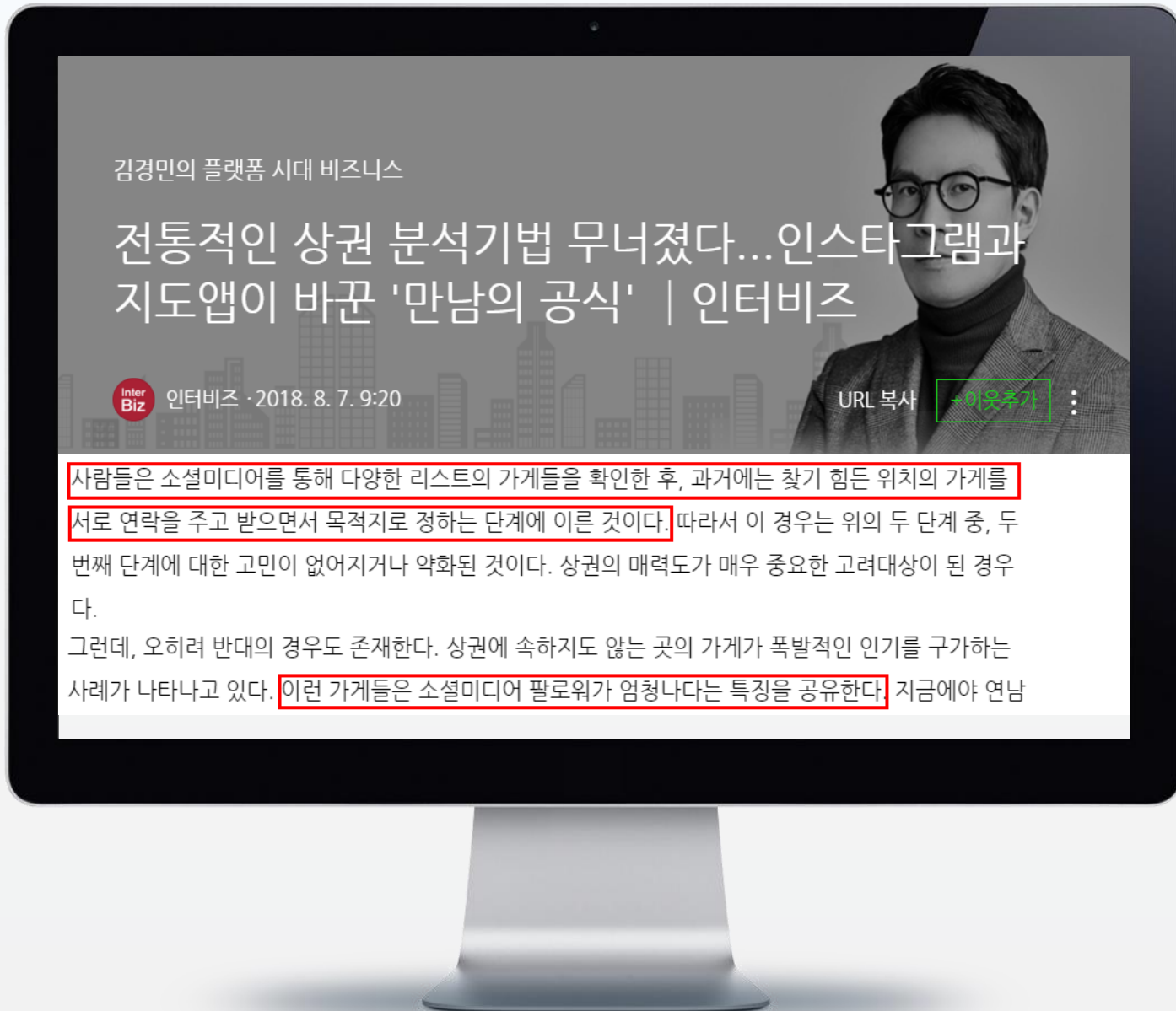
불편 감수하고도 갈 곳은 간다

사진찍기 좋은 곳이 인기

출처: 여성신문

“해시태그 상권”
“SNS 상권”
“인스타그램 상권”





김경민의 플랫폼 시대 비즈니스

전통적인 상권 분석기법 무너졌다...인스타그램과 지도앱이 바꾼 '만남의 공식' | 인터뷰

"사람들은 소셜미디어를 통해 다양한 리스트의 가게들을 확인한 후, 과거에는 찾기 힘든 위치의 가게를 서로 연락을 주고 받으면서 목적지로 정하는 단계에 이르렀다."

"상권에 속하지도 않은 곳의 가게들이 폭발적인 인기를 구가하는 사례도 있는데, 이런 가게들은 소셜 미디어 팔로워가 엄청나다라는 특징을 공유한다."

"신규 혹은 기존 고객에게 알려지지 않은 곳에 굉장히 독특한 가게가 위치하고 있다는 정보가 정보 유통 채널을 통해 전달되어야 한다. 그리고 여기에 이미지-영상 기반 소셜미디어는 대단한 역할을 수행했다."

"이 경우는 상권의 중요성이라는 첫 단계를 무시한 채 두 번째 단계인 개별 가게의 경쟁력이 주 고려대상이 된 것이다."



WHY_Instagram

WHY_Instagram #소비_트렌드 변화의 중심 #인스타그램



좋아요 191,122개

1일 전



댓글 달기...

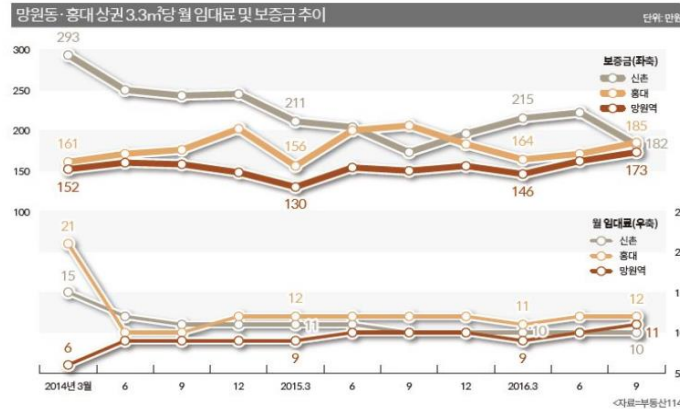


'망리단길' 아시나요...망원동 상권 '주목'

2017년 02월호

역세권 아니고 대형 쇼핑몰도 없고 간판도 없지만 입소문과 SNS로 부상
월 임대료 80% 올라도 여전히 홍대보다 저렴해 매력
투자자·창업자 꾸준히 몰리며 2000만원 권리금도 불어

상권 성공의 기본 요소로 꼽히는 역세권은 커녕 도로도 좁고 혼한 복합쇼핑몰과 유명 프랜차이즈 가게도 없다. 노후 단독주택 사이에 가게가 있고 간판도 없어 초행길이면 가게 앞에서 그 가게를 찾는 일도 흔하다. 그러나 인스타그램, 페이스북, 블로그 등 SNS(사회관계망서비스)상에서 입소문을 타고 많은 젊은이들이 평일 오후에도 이곳을 찾아 음식을 먹고 문화를 즐긴다.



망리단 상가 월세, 30개월 새 80% 올라

망원동 상권은 특색 있는 맛집들이 모이고 있다는 입소문에 월 임대료와 보증금도 오르고 있다. 부동산정보업체 부동산114에 따르면 지난해 9월 말(3분기) 기준 망원역 주변 상가 3.3㎡당 월 임대료(1층 기준)는 11만1400원 수준이다. 지난 2014년 3월 말(1분기) 기준 6만1800원에서 80% 오른 수치다. 같은 기간 홍대 지역 상가 월 임대료는 20만6700원에서 12만1500원으로 40% 가까이 빠졌다. 신촌 상권도 월 임대료가 15만2900원에서 10만3900원으로 32% 감소했다. 하지만 망원역 상권 임대료는 80%나 올랐어도 여전히 홍대보다 저렴하다.

출처: 뉴스핌, 2017.02월호 기사



WHY_Instagram

WHY_Instagram #SNS상권_임대료상승



좋아요 191,122개

1일 전

댓글 달기...

...

[인스타, 거기 어디?]문짝 하나로 SNS스타된 망원동 카페 '자판기'

[중앙일보] 입력 2017.06.26 00:01



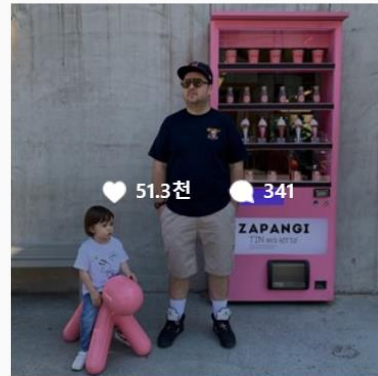
#자판기카페

게시물 18,802

팔로우

관련 해시태그 #재미도있드라 #최애캐

인기 게시물



tired_jihyunkim

tired_jihyunkim #소비_트렌드를 파악하기에 가장 유용한 SNS

soyeon_park #언급량_18,000회 이상
choi_sw 인기글은 좋아요 513,000개
carzerosushi #Business #Analysis
@tired_jihyunkim 컨셉이 분명한 카페



좋아요 191,122개

1일 전



댓글 달기...



“카페 업종의 폐업률을 아시나요? 36개월 이내에 90%예요”

“제가 여기서 2016년에 개업을 했어요.
그 때랑 지금이랑 유동인구수가 크게 차이나요.”

“과거에는 ‘인스타 셀럽’들을 통해서 인스타그램으로 홍보를 엄청 했습니다.
그런데, 지금은 하지 않아요.”

QUESTION

- ① 망원 상권은 인스타그램과 영향을 주고받는 관계에 있는가?
- ② 분석 대상에 포함된 모든 업종이 인스타그램과 영향 관계에 있는가?
아니면 영향을 받는 특정 업종이 있는가?

QUESTION

③ 연도 별로 인스타의 영향력이 달라지는가?

④ 분기 별로 세분화하여 볼 때,
인스타그램의 영향도는 계절성을 보이는가

2

가설 검증을 위한 데이터 분석

문제 해결을 위한 데이터 분석 방법론

데이터 수집

WHERE FROM 🔍

#StarTag # @ 망원 🔍 + 업종별 검색 예시

태그분석 결과 ⓘ 이용안내 '스타태그 ACCESS' 알아보기

#태그	누적게시물 수	인기글 평균 좋아요 수	인기글 평균 유지시간(h)	게시물 생산지수	반응도	트렌드지수
<input type="checkbox"/> #망원동	1,013,613	988	85.07	0.4	13	0.13
<input type="checkbox"/> #망원동카페	306,574	439	62.85	0.26	17	0.57
<input type="checkbox"/> #망원동맛집	288,902	1,482	162.97	0.14	9	0.31
<input type="checkbox"/> #망원	272,789	535	84.19	0.14	6	0.22
<input type="checkbox"/> #망원시장	88,602	1,334	879.14	0.06	3	0.31
<input type="checkbox"/> #망원카페	71,548	2,067	340.95	0.05	3	0.42
<input type="checkbox"/> #망원맛집	63,673	1,920	886.11	0.03	2	0.3
<input type="checkbox"/> #망원동티라미수	63,377	2,669	2528.43	0.02	1	0.09
<input type="checkbox"/> #망원역	49,151	682	1077.91	0.03	1	0.28
<input type="checkbox"/> #망원한강공원	37,019	634	620.02	0.03	2	0.48
<input type="checkbox"/> #망원동꽃집	27,900	187	204.57	0.02	1	0.4

HOW TO 🔍

데이터 수집 🔍

- 스타태그에서 망원과 관련된 해시태그 수집
(검색어: "망원", "망리단길", etc.)

WHERE FROM 🔍

#태그	누적게시물 수	인기글 평균 좋아요 수	인기글 평균 유지시간(h)	게시물 생산지수	반응도	트렌드지수
#망원동	1,013,613	988	85.07	0.4	13	0.13
#망원동카페	306,574	439	62.85	0.26	17	0.57
#망원동맛집	288,902	1,482	162.97	0.14	9	0.31
#망원	272,789	535	84.19	0.14	6	0.22
#망원시장	88,602	1,334	879.14	0.06	3	0.31
#망원카페	71,548	2,067	340.95	0.05	3	0.42
#망원맛집	63,673	1,920	886.11	0.03	2	0.3
#망원동티라미수	63,377	2,669	2528.43	0.02	1	0.09
#망원역	49,151	682	1077.91	0.03	1	0.28
#망원한강공원	37,019	634	620.02	0.03	2	0.48
#망원동꽃집	27,900	187	204.57	0.02	1	0.4

HOW TO 🔍

데이터 수집

- 1) 스타태그에서 망원과 관련된 해시태그 수집
(검색어: "망원", "망리단길", etc.)
- 2) 업종을 대표할 수 있는 해시태그 추출

WHERE FROM 🔍

#태그	누적게시물 수	인기글 평균 좋아요 수	인기 평균	트렌드지수
#망원동	1,013,613	988	85.07	0.13
#망원동카페	306,574	439	62.85	0.26
#망원동맛집	288,902	1,482	9	0.31
#망원	272,789	535	8	0.22
#망원시장	88,602	1,334	870.14	0.06
#망원카페	71,548	2,067	510.55	0.05
#망원맛집	63,673	1,920	2	0.3
#망원동티라미수	63,377	2,669	2528.4	0.09
#망원역	49,151	682	1077.91	0.03
#망원한강공원	37,019	634	620.02	0.03
#망원동꽃집	27,900	187	204.57	0.02

HOW TO 🔍

데이터 수집

- 스타태그에서 망원과 관련된 해시태그 수집
(검색어: "망원", "망리단길", etc.)
- 업종을 대표할 수 있는 해시태그 추출
 대분류: 서비스업
 중분류: 스포츠클럽, 미용업, 네일

 대분류: 소매업
 중분류: 의류점, 패션용품, 화초(꽃집)

 대분류: 외식업
 중분류: 한식, 중식, 양식, 일식, 패스트푸드, 제과점, 커피/음료, 호프/바

업종 분류 기준: 우리마을가게상권분석

WHERE FROM 🔍

#망원동즉석우동 #망원역맛집 #망원애견동반
 #망원동맛집 #발리인망원 #망원카페
 #망원동밥집 #망원시장맛집 #망원동카페
 #망원동칵테일 #망원동바 #망원동마카롱
 #망원동술집 #망원술집 #망원동카페추천 #망원동디저트
 #망원동와인 #망원동베이커리 #망원동꽃집
 #망원공방 #망원빵집 #망원역꽃집
 #망원동네일 #망원옷가게 #망원반영구
 #망원네일 #망원네일아트 #망원동미용실
 #망원눈썹문신

HOW TO 🔍



선정된 해시태그의 **트렌드 지수**와 **반응도 지표**를 수집 => 총 **14개 업종의 인스타그램 지표** 수집

우리마을가게상권분석서비스 <https://golmok.seoul.go.kr/>

스포츠클럽, 미용업, 네일, 의류점, 패션용품, 화초(꽃집), 한식, 중식, 양식, 일식, 패스트푸드, 제과점, 커피/음료, 호프/바

업종 현황 데이터 수집



전체 점포 수, 프랜차이즈 점포 수, 일반 점포 수,
개업 수, 폐업 수, 개업률, 폐업률,
평균 10년 영업기간, 평균 30년 영업기간,
1년 생존율, 3년 생존율, 5년 생존율

우리마을가게상권분석서비스 <https://golmok.seoul.go.kr/>

스포츠클럽, 미용업, 네일, 의류점, 패션용품, 화초(꽃집), 한식, 중식, 양식, 일식, 패스트푸드, 제과점, 커피/음료, 호프/바

업종 현황 데이터 수집



전체 점포 수, 프랜차이즈 점포 수, 일반 점포 수,
개업 수, 폐업 수, 개업률, 폐업률,
평균 10년 영업기간, 평균 30년 영업기간,
1년 생존율, 3년 생존율, 5년 생존율

```
from urllib.request import urlopen
import pandas as pd
from selenium import webdriver
from urllib.parse import quote
import requests
from bs4 import BeautifulSoup
```

웹 페이지의 HTML을 가져오는 모듈
HTML을 파싱하는 모듈

Python - 웹 크롤링

우리가게 사이트 크롤링 메인 코드 Last Checkpoint: 2019.11.30 (autosaved)

```
View Insert Cell Kernel Widgets Help Trusted
Run C Code
if k == 2: # 내분류 외식업

for i in joongbun1: # 중분류 반복문
    driver.find_element_by_xpath("//*[id='induM']/option[%d]" % i).click() # 중분류 선택
    driver.find_element_by_xpath("//*[id='infoCategory']/option[2]").click() # 카테고리 선택

    driver.find_element_by_xpath("//*[id='presentSearch']").click() # 검색 클릭
    time.sleep(1)

#####
# 크롬 드라이버 현재 페이지를 가져온 뒤 BeautifulSoup 객체로 만들
html = driver.page_source
soup = BeautifulSoup(html, 'html.parser')
# table에 있는 html 코드 가져오기
table_div = soup.find(id="table1")
tables = table_div.find_all("tr")
# 233번째 망원 1동, 234번째 망원 2동
menu_table = tables[233]
menu_table2 = tables[234]
# 텍스트만 가져오기
string1 = menu_table.get_text()
string2 = menu_table2.get_text()
split_list = re.split(' ', string1)
split_list2 = re.split(' ', string2)
```

서울시빅데이터캠퍼스 2019/11/29 (금) 상암 본사에 직접 방문

스포츠클럽, 미용업, 네일, 의류점, 패션용품, 화초(꽃집), 한식, 중식, 양식, 일식, 패스트푸드, 제과점, 커피/음료, 호프/바

업종 현황 데이터 수집



각 업종의 2017년, 2018년, 2019년
매출 데이터 수집
개인 매출 데이터를 합하여 분기 별 매출 자료 생성

1분기: 1월~3월
2분기: 4월~6월
3분기: 7월~9월
4분기: 10월~12월

서울시빅데이터캠퍼스 2019/11/29 (금) 상암 본사에 직접 방문

스포츠클럽, 미용업, 네일, 의류점, 패션용품, 화초(꽃집), 한식, 중식, 양식, 일식, 패스트푸드, 제과점, 커피/음료, 호프/바

업종 현황 데이터 수집



R - 사용 라이브러리: dplyr

각 업종의 2017년, 2018년, 2019년
매출 데이터 수집
개인 매출 데이터를 합하여 분기 별 매출 자료 생성

1분기: 1월~3월
2분기: 4월~6월
3분기: 7월~9월
4분기: 10월~12월

서울시빅데이터캠퍼스 2019/11/29 (금) 상암 본사에 직접 방문

스포츠클럽, 미용업, 네일, 의류점, 패션용품, 화초(꽃집), 한식, 중식, 양식, 일식, 패스트푸드, 제과점, 커피/음료, 호프/바

R - 사용 라이브러리: dplyr

```

1 data <- read.table("E:\\2. 내국인(집계구)_성별연령대별\\SEOUL_KOR_TOTREG_DEMO_1705.txt", sep="|", quote="", header=T, fileEncoding="utf8")
2 ourdata <- data.frame(data)
3
4 ourdata$가맹점집계구코드 <- (ourdata$'가맹점집계구코드')%/%1000000 # 블록 단위를 동 단위로 출력하기 위해 코드를 100만으로 나눈 몫
5
6 ##망원 1동 추출
7 ourdata2 <- ourdata[ourdata$가맹점집계구코드==1114069,]
8 head(ourdata2)
9 #내국인업종코드, 성별, 연령대별, 카드이용금액 추출
10 ourdata2 <- ourdata2[,c(2,8)]
11
12 ##망원 1동 추출
13 ourdata3 <- ourdata[ourdata$가맹점집계구코드==1114070,]
14 head(ourdata3)
15 #내국인업종코드, 성별, 연령대별, 카드이용금액 추출
16 ourdata3 <- ourdata3[,c(2,8)]
17
18 library(dplyr)
19 mangwon1 <- ourdata2 %>%
20   group_by(내국인업종코드) %>%
21   summarize(total = sum(카드이용금액, na.rm=TRUE))
22 mangwon2 <- ourdata3 %>%
23   group_by(내국인업종코드) %>%

```


소상공인 상권정보시스템 <http://sg.sbiz.or.kr/>

스포츠클럽, 미용업, 네일, 의류점, 패션용품, 화초(꽃집), 한식, 중식, 양식, 일식, 패스트푸드, 제과점, 커피/음료, 호프/바

업종 현황 데이터 수집



업종 별 연령 비율 데이터를 추가 수집하여,
'인스타그램 매출 변수' 데이터 조정에 이용

'상권분석' 및 '경쟁분석' 보고서

커피전문점/카페/다방	매출액	2,115	1,961	-7.28%▼	1,877	-4.28%▼	1,771	-5.65%▼	1,858	4.91%▲	1,766	-4.95%▼
	건수	1,898	1,998	5.27%▲	1,934	-3.20%▼	1,919	-0.78%▼	1,970	2.66%▲	1,887	-4.21%▼
커피점/카페	매출액	2,115	1,961	-7.28%▼	1,877	-4.28%▼	1,771	-5.65%▼	1,858	4.91%▲	1,766	-4.95%▼
	건수	1,898	1,998	5.27%▲	1,934	-3.20%▼	1,919	-0.78%▼	1,970	2.66%▲	1,887	-4.21%▼
음식	매출액	2,854	3,086	8.13%▲	2,967	-3.86%▼	2,932	-1.18%▼	3,032	3.41%▲	2,716	-10.42%▼
	건수	1,520	1,619	6.51%▲	1,569	-3.09%▼	1,587	1.15%▲	1,654	4.22%▲	1,466	-11.37%▼

· 성별/연령대별 매출 비율

(단위 : 만원)

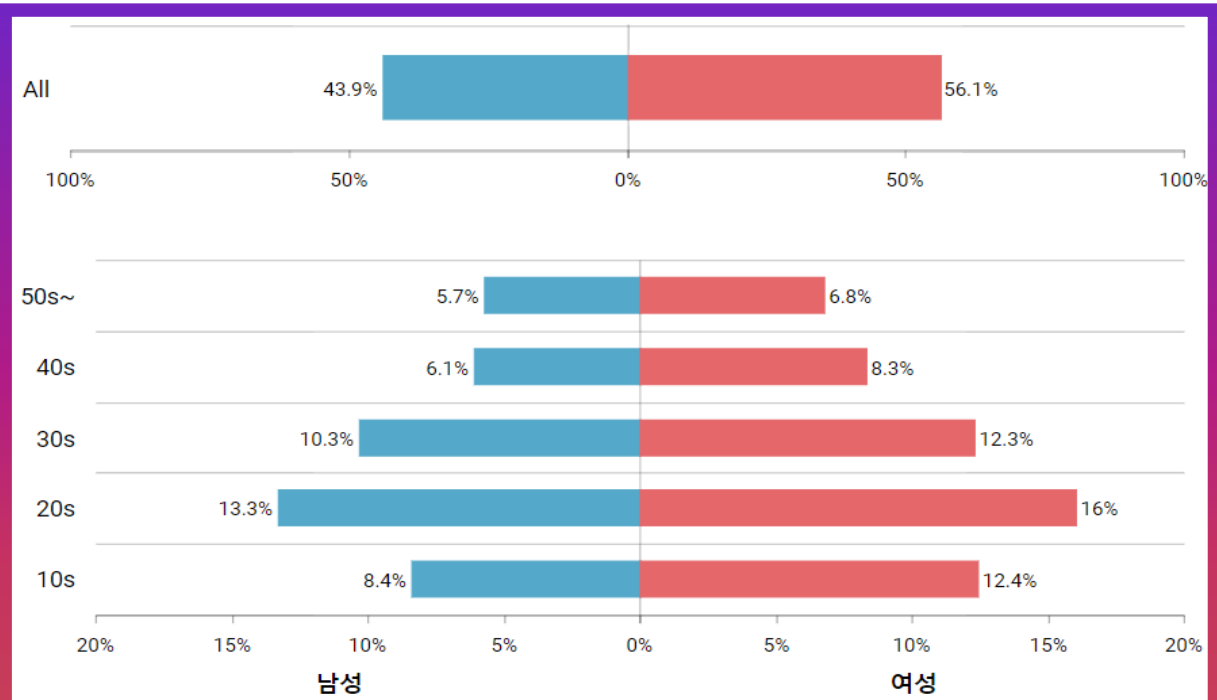
업종	구분	성별		연령별					
		남성	여성	10대	20대	30대	40대	50대	60대이상
제1선택영역	매출액	592	1,173	0	515	753	258	187	52
	비율	33.5%	66.5%	0.0%	29.2%	42.7%	14.6%	10.6%	2.9%

데이터 베이스 구축 Excel - VLOOKUP 함수

dong	dae	joong	year	bungi	trendscore	react	averSales	totalStores	FranS tores	General Stores	Opening Num	Closing Num	Openin gRatio	Closing Ratio	last10Ave ragePerio	last30Aver agePeriod
mangwon1	service	sportsClub	2017	1	0.15	0	5032130.25	8	0	8	1	0	12.5	0	2.1	5.5
mangwon2	service	sportsClub	2017	1	0.15	0	42876716.3	4	0	4	0	1	0	25	2.9	7.4
mangwon1	service	hairDressing	2017	1	0.5	0	2800921.71	51	0	51	1	0	2	0	4.1	5.9
mangwon2	service	hairDressing	2017	1	0.5	0	1419785.76	62	1	61	1	0	1.6	0	3.9	6.3
mangwon1	service	nailShop	2017	1	1.1	0.3	627908.571	7	0	7	0	0	0	0	2.9	3.7
mangwon2	service	nailShop	2017	1	1.1	0.3	954188.8	5	0	5	0	0	0	0	2.6	3.5
mangwon1	service	sportsClub	2017	2	0.1	0	7506560.43	7	0	7	0	1	0	14.3	2.2	5.5
mangwon2	service	sportsClub	2017	2	0.1	0	42043226.5	4	0	4	0	0	0	0	1.7	7.5
mangwon1	service	hairDressing	2017	2	0.8	0	3208648.63	51	0	51	0	0	0	0	4.1	6
mangwon2	service	hairDressing	2017	2	0.8	0	816203.524	63	1	62	2	1	3.2	1.6	3.6	6.3
mangwon1	service	nailShop	2017	2	1.5	0.4	631756.5	8	0	8	1	0	12.5	0	2.9	3.6
mangwon2	service	nailShop	2017	2	1.5	0.4	6163761.5	6	0	6	1	0	16.7	0	2.4	3.3
mangwon1	service	sportsClub	2017	3	0.2	0	7650891.17	6	0	6	0	1	0	16.7	2.3	5.6
mangwon2	service	sportsClub	2017	3	0.2	0	53753452.3	4	0	4	0	0	0	0	1.9	7.7
mangwon1	service	hairDressing	2017	3	1.63	0.03	3078995.7	50	0	50	0	1	0	2	4.2	6.1
mangwon2	service	hairDressing	2017	3	1.63	0.03	1531157.89	63	1	62	2	2	3.2	3.2	3.7	6.3
mangwon1	service	nailShop	2017	3	1	0.5	1158753.88	8	0	8	0	0	0	0	3	3.8
mangwon2	service	nailShop	2017	3	1	0.5	494713.833	6	0	6	0	0	0	0	2.6	3.4
mangwon1	service	sportsClub	2017	4	0.4	0	6251133.67	6	0	6	1	1	16.7	16.7	2.1	5.3
mangwon2	service	sportsClub	2017	4	0.4	0	60705422	4	0	4	0	0	0	0	2.1	7.9
mangwon1	service	hairDressing	2017	4	1.97	0.07	2132960.2	51	0	51	2	1	3.9	2	4	6.1
mangwon2	service	hairDressing	2017	4	1.97	0.07	1424357.24	63	1	62	1	1	1.6	1.6	3.8	6.3
mangwon1	service	nailShop	2017	4	1.07	0.5	1081591.63	8	0	8	0	0	0	0	3.1	3.9
mangwon2	service	nailShop	2017	4	1.07	0.5	468645.167	6	0	6	0	0	0	0	2.7	3.6
mangwon1	service	sportsClub	2018	1	0.45	0	27075680.7	6	0	6	0	0	0	0	2.2	5.4
mangwon2	service	sportsClub	2018	1	0.45	0	55474377.8	5	0	5	1	0	20	0	1.7	6.9
mangwon1	service	hairDressing	2018	1	1.57	0.07	2826079.21	52	0	52	1	0	1.9	0	4	6.2
mangwon2	service	hairDressing	2018	1	1.57	0.07	1294316.85	67	1	66	5	1	7.5	1.5	3.5	6.2
mangwon1	service	nailShop	2018	1	1.1	0.6	529940.375	8	0	8	0	0	0	0	3.3	4
mangwon2	service	nailShop	2018	1	1.1	0.6	679705.333	6	0	6	0	0	0	0	2.9	3.7
mangwon1	service	sportsClub	2018	2	0.45	0	9713157	6	0	6	0	0	0	0	1.5	5.5
mangwon2	service	sportsClub	2018	2	0.45	0	45696397.4	5	0	5	0	0	0	0	1.9	7.1
mangwon1	service	hairDressing	2018	2	2.2	0.1	7659403.72	50	0	50	1	3	2	6	4	6.2

데이터 전처리

1. 총 매출 변수를 전체 점포수로 나누어 점포당 평균 매출 구한 후, 각 업종 별 20,30대의 비율을 곱하여 '인스타그램 표적 매출' 변수 생성



- 인스타그램 활성 사용자 기준, 20대 여성이 가장 활발
- 전반적으로는 20대~30대의 비율이 압도적으로 높고, 성별 비율의 차이는 유의미하게 벌어지지 않음

- 인스타그램의 영향력을 보다 정확하게 파악하기 위해서 변수를 가공할 필요가 있음
- 인스타그램의 주요 사용자가 10~30대인 것을 감안하여, 앞서 수집한 매출, 거래 건수 등의 데이터를 가공할 것
- 가공 방법: 업종 별 연령대의 비율을 매출, 거래 건수 등의 데이터에 적용하여 보다 신뢰성 있는 변수 생성

2. Data Rescaling (정규화 실시) - 데이터의 범위를 0과 1로 변환.
 데이터 분포를 조정변수 간의 측정 단위와 데이터의 스케일이 다르기 때문임

=((H2 - MIN(\$H\$2:\$H\$281)) / (MAX(\$H\$2:\$H\$281) - MIN(\$H\$2:\$H\$281)))							
D	E	F	G	H	I	J	K
year	bungi	trendscore	react	averSales	jungggyuSa	totalStore	FranSto
2017	1	0.15	0	5032130	0.060008	8	
2017	1	0.15	0	42876716	0.517688	4	
2017	1	0.5	0	2800922	0.033025	51	

3. 연속형 변수들끼리 상관 분석 실시 => 서로 상관성이 높은 변수들 제거 => 모델의 혼란과 다중공산성 제거

```
> rcorr(as.matrix(ourdata2), type="pearson")
```

	트렌드지수	반응도	매출	전체점포수	프랜차이즈점포수	일반점포수	개업수	폐업수	개업률	폐업률	최근10년평균영업기간	최근30년평균영업기간	
트렌드지수	1.00	0.21	0.11	-0.06	-0.04	-0.06	0.03	-0.07	0.14	-0.02	0.13	-0.07	
반응도	0.21	1.00	-0.03	0.41	0.42	0.38	0.52	0.42	0.17	0.09	-0.20	-0.37	
매출	0.11	-0.03	1.00	-0.26	-0.16	-0.26	-0.17	-0.15	-0.05	0.04	0.01	0.03	
전체점포수	-0.06	0.41	-0.26	1.00	0.46	1.00	0.72	0.73	0.04	0.01	0.05	-0.05	
프랜차이즈점포수	-0.04	0.42	-0.16	0.46	1.00	0.39	0.33	0.42	0.02	0.10	-0.03	-0.05	
일반점포수	-0.06	0.38	-0.26	1.00	0.39	1.00	0.72	0.73	0.04	0.00	0.06	-0.05	
개업수	0.03	0.52	-0.17	0.72	0.33	0.72	1.00	0.65	0.50	0.09	-0.23	-0.31	
폐업수	-0.07	0.42	-0.15	0.73	0.42	0.73	0.65	1.00	0.12	0.43	-0.14	-0.20	
개업률	0.14	0.17	-0.05	0.04	0.02	0.04	0.50	0.12	1.00	0.10	-0.35	-0.38	
폐업률	-0.02	0.09	0.04	0.01	0.10	0.00	0.09	0.43	0.10	1.00	-0.21	-0.15	
최근10년평균영업기간	0.13	-0.20	0.01	0.05	-0.03	0.06	-0.23	-0.14	-0.35	-0.21	1.00	0.63	
최근30년평균영업기간	-0.07	-0.37	0.03	-0.05	-0.05	-0.05	-0.31	-0.20	-0.38	-0.15	0.63	1.00	

n= 280

P

	트렌드지수	반응도	매출	전체점포수	프랜차이즈점포수	일반점포수	개업수	폐업수	개업률	폐업률	최근10년평균영업기간	최근30년평균영업기간	
트렌드지수		0.0003	0.0786	0.3425	0.5562	0.3506	0.6668	0.2335	0.0167	0.7054	0.0329	0.2709	
반응도	0.0003		0.5890	0.0000	0.0000	0.0000	0.0000	0.0000	0.0044	0.1415	0.0010	0.0000	
매출	0.0786	0.5890		0.0000	0.0075	0.0000	0.0036	0.0113	0.4335	0.5595	0.8905	0.6214	
전체점포수	0.3425	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.5555	0.8456	0.3970	0.4044	
프랜차이즈점포수	0.5562	0.0000	0.0075	0.0000		0.0000	0.0000	0.0000	0.8021	0.0951	0.6505	0.3962	
일반점포수	0.3506	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.5553	0.9555	0.3581	0.4291	
개업수	0.6668	0.0000	0.0036	0.0000	0.0000	0.0000		0.0000	0.0000	0.1212	0.0001	0.0000	
폐업수	0.2335	0.0000	0.0113	0.0000	0.0000	0.0000	0.0000		0.0546	0.0000	0.0158	0.0010	
개업률	0.0167	0.0044	0.4335	0.5555	0.8021	0.5553	0.0000	0.0546		0.0941	0.0000	0.0000	
폐업률	0.7054	0.1415	0.5595	0.8456	0.0951	0.9555	0.1212	0.0000	0.0941		0.0004	0.0101	
최근10년평균영업기간	0.0329	0.0010	0.8905	0.3970	0.6505	0.3581	0.0001	0.0158	0.0000	0.0004		0.0000	
최근30년평균영업기간	0.2709	0.0000	0.6214	0.4044	0.3962	0.4291	0.0000	0.0010	0.0000	0.0101	0.0000		

3. 연속형 변수들끼리 상관 분석 실시 => 서로 상관성이 높은 변수들 제거
=> 모델의 혼란과 다중공산성 제거

```
> rcorr(as.matrix(ourdata2), type="pearson")
```

	트렌드지수	반응도	매출	전체점포수	프랜차이즈점포수	일반점포수	개업수	폐업수	개업률	폐업률	최근10년평균영업기간	최근30년평균영업기간
트렌드지수	1.00	0.21	0.11	-0.06	-0.04	-0.06	0.03	-0.07	0.14	-0.02	0.13	-0.07
반응도	0.21	1.00	0.03	0.41	0.42	0.38	0.52	0.43	0.17	0.09	-0.30	-0.37
매출	0.11	0.03	1.00	-0.26	-0.16	-0.26	0.72	0.73	0.04	0.00	0.05	0.03
전체점포수	-0.06	0.41	-0.26	1.00	0.46	0.39	1.00	0.72	0.73	0.04	0.06	-0.05
프랜차이즈점포수	-0.04	0.42	-0.16	0.46	1.00	0.39	1.00	0.72	0.73	0.04	0.06	-0.05
일반점포수	-0.06	0.38	-0.26	0.39	0.39	1.00	0.72	0.73	0.04	0.00	0.06	-0.05
개업수	0.03	0.52	0.72	0.72	0.72	1.00	0.72	0.73	0.04	0.00	0.06	-0.05
폐업수	-0.07	0.43	0.73	0.73	0.73	0.73	1.00	0.72	0.73	0.04	0.06	-0.05
개업률	0.14	0.17	-0.05	0.04	0.04	0.04	0.04	0.09	1.00	0.10	-0.35	-0.38
폐업률	-0.02	0.09	0.04	0.01	0.01	0.01	0.01	0.09	0.10	1.00	-0.21	-0.15
최근10년평균영업기간	0.13	-0.20	0.01	0.05	-0.03	0.06	-0.23	-0.14	-0.35	-0.21	1.00	0.63
최근30년평균영업기간	-0.07	-0.37	0.03	-0.05	-0.05	-0.05	-0.05	-0.05	-0.38	-0.15	0.63	1.00

- ✓ 일반 점포 수와 개업 수, 폐업 수의 상관계수가 각각 0.72, 0.73으로 나왔고, p-value도 유효한 수치
=> 개업 수, 폐업 수를 제거, 개업률과 폐업률 사용
- ✓ 최근 10년 평균 영업 기간과 최근 30년 평균 영업 기간의 상관계수가 0.63, p-value가 0.00으로 유효함
=> 최근 10년 평균 영업 기간만 사용
- ✓ 전체 점포 수 변수 제거 + 일반 점포 수와 프랜차이즈 점포 수 변수만 사용
- ✓ 1년 생존율, 3년 생존율, 5년 생존율은 결측치가 많아서 사용하지 않기로 결정

```
P
```

	트렌드지수	반응도	매출	전체점포수	프랜차이즈점포수	일반점포수	개업수	폐업수	개업률	폐업률	최근10년평균영업기간	최근30년평균영업기간
트렌드지수		0.0003	0.0786	0.3425	0.5562	0.3506	0.6668	0.2335	0.0167	0.7054	0.0329	0.2709
반응도	0.0003		0.5890	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0044	0.1415	0.0010
매출	0.0786	0.5890		0.0000	0.0075	0.0000	0.0000	0.0000	0.0000	0.0113	0.4335	0.5595
전체점포수	0.3425	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
프랜차이즈점포수	0.5562	0.0000	0.0075	0.0000		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
일반점포수	0.3506	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
개업수	0.6668	0.0000	0.0036	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000	0.0000
폐업수	0.2335	0.0000	0.0113	0.0000	0.0000	0.0000	0.0000		0.0000	0.0000	0.0000	0.0000
개업률	0.0167	0.0044	0.4335	0.5555	0.8021	0.5553	0.0000	0.0546		0.0941	0.0000	0.0000
폐업률	0.7054	0.1415	0.5595	0.8456	0.0951	0.9555	0.1212	0.0000	0.0941		0.0004	0.0101
최근10년평균영업기간	0.0329	0.0010	0.8905	0.3970	0.6505	0.3581	0.0001	0.0158	0.0000	0.0004		0.0000
최근30년평균영업기간	0.2709	0.0000	0.6214	0.4044	0.3962	0.4291	0.0000	0.0010	0.0000	0.0101	0.0000	

최종적으로 매출, 프랜차이즈 점포 수, 일반 점포 수, 개업률, 폐업률, 최근 10년 평균 영업 기간을
업종과 관련된 지표로 사용하기로 결정함

4. 명목변수 Dummy 변수화

업종의 reference는 중식, 분기의 reference는 4분기, 년도의 reference는 2019년으로 설정

R - fastDummies 라이브러리 사용 업종, 년도, 분기 변수를 더미 변수화

```
install.packages("fastDummies") #D를 꼭 대문자로 써야 함
library("fastDummies")
## 더미 변수호
df<-dummy_cols(junggyu, select_columns=c("joong","year",'bunggi'))
#head(df)
str(df)
```

[illegible]

최종 데이터 베이스

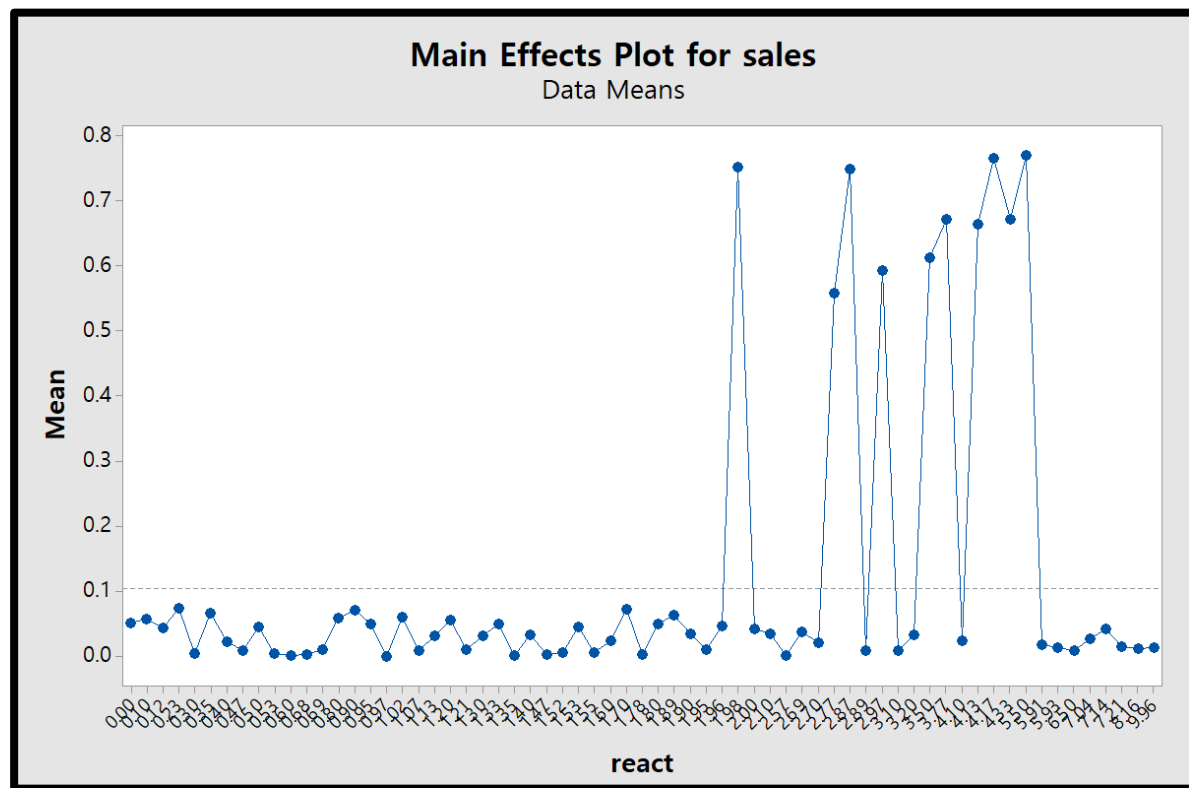
280 obs. Of 35 variables

dong	dae	joong	year	bungi	trendscore	react	sales	averSales	FranStores	GeneralSt	OpeningR	ClosingRa	last10Aver	joong_bak	joong_chir	joong_clot	joong_coff	joong_fast	joong_fast	joong_flow	joong_hai	joong_hop	joong_jap	joong_kor		
	1	service	sportsClub	2017	1	0.15	0	0.014637	0.060008	0	0.042857	0.437063	0	0.3	0	0	0	0	0	0	0	0	0	0		
	0	service	sportsClub	2017	1	0.15	0	0.064016	0.517688	0	0.014286	0	0.874126	0.566667	0	0	0	0	0	0	0	0	0	0		
	1	service	hairDressir	2017	1	0.5	0	0.053233	0.033025	0	0.35	0.06993	0	0.966667	0	0	0	0	0	0	0	1	0	0	0	
	0	service	hairDressir	2017	1	0.5	0	0.032609	0.016322	0.1	0.421429	0.055944	0	0.9	0	0	0	0	0	0	0	1	0	0	0	
	1	service	nailShop	2017	1	1.1	0.3	0.001146	0.006745	0	0.035714	0	0	0.566667	0	0	0	0	0	0	0	0	0	0	0	
	0	service	nailShop	2017	1	1.1	0.3	0.001287	0.010691	0	0.021429	0	0	0.466667	0	0	0	0	0	0	0	0	0	0	0	0
	1	service	sportsClub	2017	2	0.1	0	0.019261	0.089933	0	0.035714	0	0.5	0.333333	0	0	0	0	0	0	0	0	0	0	0	
	0	service	sportsClub	2017	2	0.1	0	0.062761	0.507609	0	0.014286	0	0	0.166667	0	0	0	0	0	0	0	0	0	0	0	0
	1	service	hairDressir	2017	2	0.8	0	0.061056	0.037956	0	0.35	0	0	0.966667	0	0	0	0	0	0	0	1	0	0	0	0
	0	service	hairDressir	2017	2	0.8	0	0.018837	0.009022	0.1	0.428571	0.111888	0.055944	0.8	0	0	0	0	0	0	0	1	0	0	0	0
	1	service	nailShop	2017	2	1.5	0.4	0.001394	0.006791	0	0.042857	0.437063	0	0.566667	0	0	0	0	0	0	0	0	0	0	0	0
	0	service	nailShop	2017	2	1.5	0.4	0.013406	0.073694	0	0.028571	0.583916	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0
	1	service	sportsClub	2017	3	0.2	0	0.016762	0.091679	0	0.028571	0	0.583916	0.366667	0	0	0	0	0	0	0	0	0	0	0	0
	0	service	sportsClub	2017	3	0.2	0	0.080384	0.649228	0	0.014286	0	0	0.233333	0	0	0	0	0	0	0	0	0	0	0	0
	1	service	hairDressir	2017	3	1.63	0.03	0.05741	0.036388	0	0.342857	0	0.06993	1	0	0	0	0	0	0	0	1	0	0	0	0
	0	service	hairDressir	2017	3	1.63	0.03	0.035783	0.017668	0.1	0.428571	0.111888	0.111888	0.833333	0	0	0	0	0	0	0	1	0	0	0	0
	1	service	nailShop	2017	3	1	0.5	0.00298	0.013165	0	0.042857	0	0	0.6	0	0	0	0	0	0	0	0	0	0	0	0
	0	service	nailShop	2017	3	1	0.5	0.000609	0.005134	0	0.028571	0	0	0.466667	0	0	0	0	0	0	0	0	0	0	0	0
	1	service	sportsClub	2017	4	0.4	0	0.013603	0.07475	0	0.028571	0.583916	0.583916	0.3	0	0	0	0	0	0	0	0	0	0	0	0
	0	service	sportsClub	2017	4	0.4	0	0.090845	0.733303	0	0.014286	0	0	0.3	0	0	0	0	0	0	0	0	0	0	0	0
1	service	hairDressir	2017	4	1.97	0.07	0.040417	0.024947	0	0.35	0.136364	0.06993	0.933333	0	0	0	0	0	0	0	1	0	0	0	0	
0	service	hairDressir	2017	4	1.97	0.07	0.033252	0.016377	0.1	0.428571	0.055944	0.055944	0.866667	0	0	0	0	0	0	0	1	0	0	0	0	
1	service	nailShop	2017	4	1.07	0.5	0.002747	0.012232	0	0.042857	0	0	0.633333	0	0	0	0	0	0	0	0	0	0	0	0	
0	service	nailShop	2017	4	1.07	0.5	0.00055	0.004819	0	0.028571	0	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	

데이터 분석&시사점

ANOVA 분산분석

반응도와 매출의 상관분석을 실시한 결과, 상관계수는 작게 나왔으나, 분산 분석을 실시했을 때는 유의한 결과를 도출할 수 있었음.



✓ 분석 결과를 시각화 하여 그래프로 확인
 ⇒ 반응도가 증가하면 매출이 증가하는 양상
 => 예외도 존재

✓ 반응도는 매출에 영향을 미치지만,
 모든 경우에서 항상 그러한 것은 아님

✓ 반응도가 상권 활성화에 영향을 어느
 정도 미치고 있다는 가정을 세울 수는 있음

반응도 상승 -> 매출 상승

① 망원의 업종은 인스타그램과 영향을
주고받는 관계에 있는가?

ANOVA 분산분석

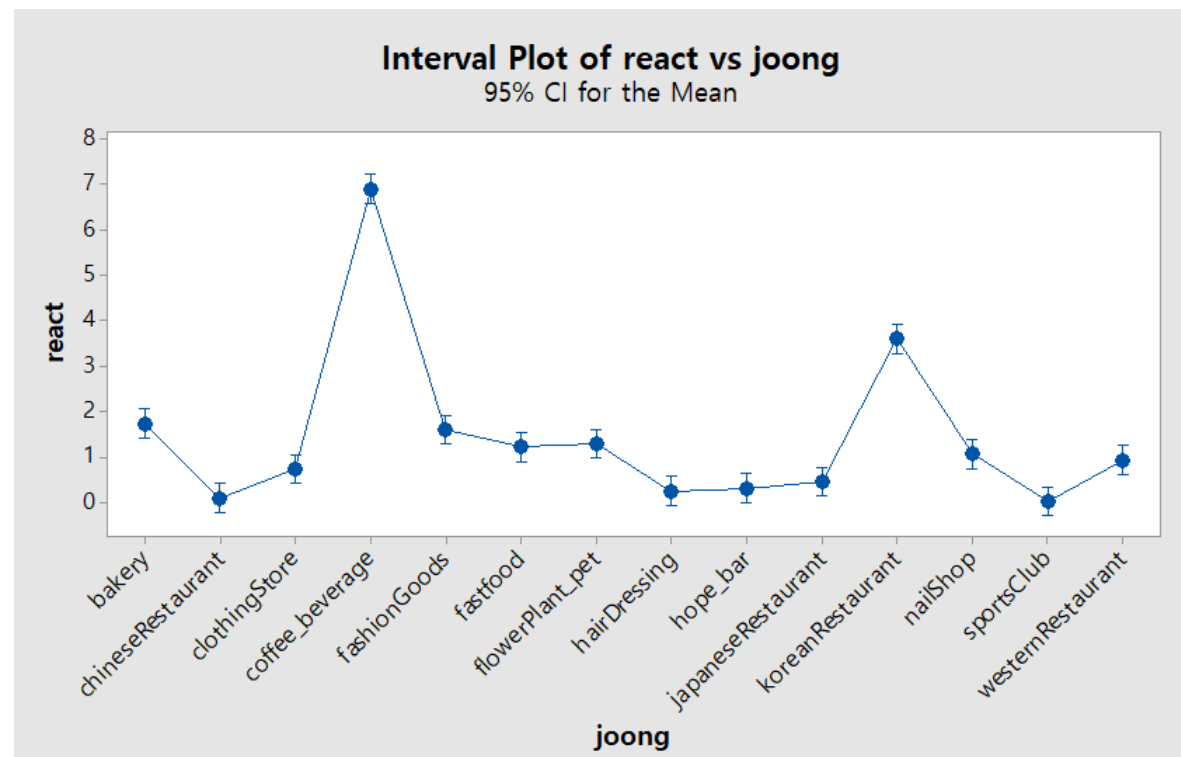
ANOVA: react versus joong

Factor Type Levels
joong fixed 14

Factor Values
joong bakery, chineseRestaurant, clothingStore, coffee_beverag

Analysis of Variance for react

Source	DF	SS	MS	F	P
joong	13	859.247	66.096	126.63	0.000
Error	266	138.845	0.522		
Total	279	998.092			



업종별 차이에 따른 반응도의 차이가 있다는 결과

다중회귀분석

종속변수: 트렌드 지수, 독립변수: 각 업종

Residuals:

Min	1Q	Median	3Q	Max
-1.9780	-0.5345	-0.1275	0.3498	3.7520

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	0.3940	0.2318	1.700	0.090308	.
joong_bakery	2.4550	0.3278	7.490	1.02e-12	***
joong_clothingStore	0.9290	0.3278	2.834	0.004945	**
joong_coffee_beverage	1.2920	0.3278	3.942	0.000103	***
joong_fashionGoods	1.2190	0.3278	3.719	0.000244	***
joong_fastfood	1.1265	0.3278	3.437	0.000683	***
joong_flowerPlant_pet	2.6400	0.3278	8.054	2.70e-14	***
joong_hairDressing	2.0840	0.3278	6.358	8.85e-10	***
joong_hope_bar	0.9910	0.3278	3.023	0.002743	**
joong_japaneseRestaurant	0.5000	0.3278	1.525	0.128336	
joong_koreanRestaurant	0.4210	0.3278	1.284	0.200111	
joong_nailShop	0.9080	0.3278	2.770	0.005997	**
joong_sportsClub	-0.0630	0.3278	-0.192	0.847727	
joong_westernRestaurant	2.1160	0.3278	6.456	5.08e-10	***
joong_chineseRestaurant	NA	NA	NA	NA	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.037 on 266 degrees of freedom

Multiple R-squared: 0.4022, Adjusted R-squared: 0.373

F-statistic: 13.77 on 13 and 266 DF, p-value: < 2.2e-16



R-squared 값이 0.4대로 낮게 나옴

다중회귀분석

종속변수: 반응도, 독립변수: 각 업종

```
Call:
lm(formula = react ~ joong_bakery + joong_clothingStore + joong_coffee_beverage +
    joong_fashionGoods + joong_fastfood + joong_flowerPlant_pet +
    joong_hairDressing + joong_hope_bar + joong_japaneseRestaurant +
    joong_koreanRestaurant + joong_nailShop + joong_sportsClub +
    joong_westernRestaurant + joong_chineseRestaurant, data = our)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-2.7990 -0.4290 -0.0230  0.2552  3.0610
```

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.1020	0.1615	0.631	0.528334
joong_bakery	1.6330	0.2285	7.148	8.51e-12 ***
joong_clothingStore	0.6400	0.2285	2.801	0.005464 **
joong_coffee_beverage	6.7970	0.2285	29.750	< 2e-16 ***
joong_fashionGoods	1.5010	0.2285	6.570	2.64e-10 ***
joong_fastfood	1.1170	0.2285	4.889	1.75e-06 ***
joong_flowerPlant_pet	1.1960	0.2285	5.235	3.35e-07 ***
joong_hairDressing	0.1550	0.2285	0.678	0.498086
joong_hope_bar	0.2100	0.2285	0.919	0.358840
joong_japaneseRestaurant	0.3570	0.2285	1.563	0.119339
joong_koreanRestaurant	3.4970	0.2285	15.306	< 2e-16 ***
joong_nailShop	0.9670	0.2285	4.233	3.18e-05 ***
joong_sportsClub	-0.0790	0.2285	-0.346	0.729780
joong_westernRestaurant	0.8270	0.2285	3.620	0.000353 ***
joong_chineseRestaurant	NA	NA	NA	NA

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7225 on 266 degrees of freedom
Multiple R-squared: 0.8609, Adjusted R-squared: 0.8541
F-statistic: 126.6 on 13 and 266 DF, p-value: < 2.2e-16

R-squared 값이 0.86대로 나옴
: 준수한 값

반응도가 트렌드 지수에 비해 인플루언서의 영향을 많이 반영하기 때문에 이런 결과가 나온 것으로 추측됨

다중회귀분석

종속변수: 반응도, 독립변수: 각 업종

```
Call:
lm(formula = react ~ joong_bakery + joong_clothingStore + joong_coffee_beverage +
    joong_fashionGoods + joong_fastfood + joong_flowerPlant_pet +
    joong_hairDressing + joong_hope_bar + joong_japaneseRestaurant +
    joong_koreanRestaurant + joong_nailShop + joong_sportsClub +
    joong_westernRestaurant + joong_chineseRestaurant, data = our)
```

Residuals:

Min	1Q	Median	3Q	Max
-2.7990	-0.4290	-0.0230	0.2552	3.0610

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.1020	0.1615	0.631	0.528334
joong_bakery	1.6330	0.2285	7.148	8.51e-12 ***
joong_clothingStore	0.6400	0.2285	2.801	0.005464 **
joong_coffee_beverage	6.7970	0.2285	29.750	< 2e-16 ***
joong_fashionGoods	1.5010	0.2285	6.570	2.64e-10 ***
joong_fastfood	1.1170	0.2285	4.889	1.75e-06 ***
joong_flowerPlant_pet	1.1960	0.2285	5.235	3.35e-07 ***
joong_hairDressing	0.1550	0.2285	0.678	0.498086
joong_hope_bar	0.2100	0.2285	0.919	0.358840
joong_japaneseRestaurant	0.3570	0.2285	1.563	0.119339
joong_koreanRestaurant	3.4970	0.2285	15.306	< 2e-16 ***
joong_nailShop	0.9670	0.2285	4.233	3.18e-05 ***
joong_sportsClub	-0.0790	0.2285	-0.346	0.729780
joong_westernRestaurant	0.8270	0.2285	3.620	0.000353 ***
joong_chineseRestaurant	NA	NA	NA	NA

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7225 on 266 degrees of freedom
 Multiple R-squared: 0.8609, Adjusted R-squared: 0.8541
 F-statistic: 126.6 on 13 and 266 DF, p-value: < 2.2e-16



종속변수: 반응도

망원의 업종은 인스타그램과 영향을 주고받는 관계에 있는가?

망원의 특정 업종은 인스타그램과 영향을 주고받는 관계에 있다. 이때, '트렌드 지수'와 '반응도' 중, '반응도'가 인스타그램과의 관계를 확인하기 좋은 지표라는 결과를 도출하였다.

트렌드 지수

생성 게시물 수 대비 반응도에 대한 값
→ 게시물 수가 적는데 반응도가 높으면 '트렌디'하다고 판단



연관성 적음

반응도

해당 해시태그가 쓰인 게시물에 대한 2차 반응
(좋아요, 댓글, 사진/영상 확인, 게시물 노출 정도)



연관성 높음

망원의 업종은 인스타그램과 영향을 주고받는 관계에 있는가?

망원의 특정 업종은 인스타그램과 영향을 주고받는 관계에 있다. 이때, '트렌드 지수'와 '반응도' 중, '반응도'가 인스타그램과의 관계를 확인하기 좋은 지표라는 결과를 도출하였다.

WHY 🔍

트렌드 지수

생성 게시물 수 대비 반응도에 대한 값
→ 게시물 수가 적는데 반응도가 높으면 '트렌디'하다고 판단

✓ 절대적 게시물 수가 적어야 트렌드 지수가 높기 때문. 즉, 트렌드 지수가 높은 해시태그가 쓰인 게시물은 노출 수가 적음

반응도

해당 해시태그가 쓰인 게시물에 대한 2차 반응
(좋아요, 댓글, 사진/영상 확인, 게시물 노출 정도)

✓ 게시물 수가 적는데 반응도가 높은 해시태그는 홍보용/광고용인 것들이 다수이다. 따라서 소비자의 수요나 행동을 설명하기 어렵다.

✓ 반응도는 잠재적 소비자의 직접적인 관심을 나타내는 지표이기 때문에 연관성이 높다.

➡ **해시태그가 사용된 게시물에 사람들이 '얼마나 반응'하였는지가 업종을 설명할 때 중요한 지표가 된다**

② 분석 대상에 포함된 모든 업종이
인스타그램과 영향 관계에 있는가?
아니면 영향을 받는 특정 업종이 있는가?

Stepwise 전진선택법

최종 결과

```
Call:
lm(formula = react ~ joong_coffee_beverage + joong_koreanRestaurant +
    joong_bakery + joong_fashionGoods + joong_flowerPlant_pet +
    joong_fastfood + joong_nailShop + joong_westernRestaurant +
    joong_clothingStore + joong_japaneseRestaurant, data = our)
```

```
Residuals:
    Min       1Q   Median       3Q      Max
-2.79900 -0.42900 -0.08125  0.30012  3.06100
```

```
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.17350    0.08064   2.152  0.0323 *
joong_coffee_beverage  6.72550    0.18031  37.301 < 2e-16 ***
joong_koreanRestaurant  3.42550    0.18031  18.998 < 2e-16 ***
joong_bakery       1.56150    0.18031   8.660 4.43e-16 ***
joong_fashionGoods  1.42950    0.18031   7.928 5.96e-14 ***
joong_flowerPlant_pet  1.12450    0.18031   6.237 1.73e-09 ***
joong_fastfood     1.04550    0.18031   5.798 1.87e-08 ***
joong_nailShop      0.89550    0.18031   4.967 1.21e-06 ***
joong_westernRestaurant 0.75550    0.18031   4.190 3.78e-05 ***
joong_clothingStore  0.56850    0.18031   3.153  0.0018 **
joong_japaneseRestaurant 0.28550    0.18031   1.583  0.1145
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 0.7212 on 269 degrees of freedom
Multiple R-squared:  0.8598,    Adjusted R-squared:  0.8546
F-statistic: 165 on 10 and 269 DF,  p-value: < 2.2e-16
```



**‘커피/음료, 한식, 제과점,
화초/애완, 패스트푸드, 양식,
의류점, 패션용품, 네일’**

**위 업종들이 인스타그램 지표 중
반응도 생성에 영향을 주고 있다는
결론 도출**

다중공산성 검사

```
> vif(ggg)
joong_coffee_beverage    joong_koreanRestaurant    joong_bakery    joong_fashionGoods    joong_sportsClub    joong_chineseRestaurant
               1.160714                1.160714                1.160714                1.160714                1.160714                1.160714
joong_hairDressing    joong_hope_bar    joong_japaneseRestaurant    joong_clothingStore
               1.160714                1.160714                1.160714                1.160714
```



VIF 값이 10 이상인 변수가 없었음. 즉, **다중공산성이 없다**는 것을 알 수 있음

표준화 계수 구하기

```
> lm.beta(gg)

Call:
lm(formula = react ~ joong_coffee_beverage + joong_koreanRestaurant +
    joong_bakery + joong_fashionGoods + joong_flowerPlant_pet +
    joong_fastfood + joong_nailShop + joong_westernRestaurant +
    joong_clothingStore, data = our)

Standardized Coefficients::
      (Intercept) joong_coffee_beverage joong_koreanRestaurant joong_bakery joong_fashionGoods joong_flowerPlant_pet joong_fastfood
      0.00000000      0.90961782      0.45947404      0.20521100      0.18720525      0.14560105      0.13482488
joong_nailShop joong_westernRestaurant joong_clothingStore
      0.11436380      0.09526679      0.06975865
```



업종 간의 인스타그램 지표 생성에
대한 영향력 비교 가능

- 1등: 커피/음료 (0.91)
- 2등: 한식 (0.46)
- 3등: 제과점 (0.21)
- 4등: 패션용품 (0.19)
- 5등: 화초 (0.15)
- 6등: 패스트푸드 (0.13)
- 7등: 네일샵 (0.11)
- 8등: 의류점 (0.07)
- 9등: 양식 (0.01)

② 분석 대상에 포함된 모든 업종이 인스타그램과 영향 관계에 있는가?

분석 대상의 모든 업종들이 인스타그램과 영향 관계에 있는 것은 아닌 것으로 분석되었다. 즉, 업종 별로 인스타그램과의 영향 정도는 다르다.

연관이 있는 업종

외식업

: 커피/음료, 한식, 제과점, 양식, 패스트푸드

소매업

: 패션용품, 의류점, 화초(꽃집)

서비스업

: 네일

연관이 없는 업종

외식업

: 중식, 일식, 호프/바

소매업

: -

서비스업

: 스포츠클럽, 미용실

② 분석 대상에 포함된 모든 업종이 인스타그램과 영향 관계에 있는가?

분석 대상의 모든 업종들이 인스타그램과 영향 관계에 있는 것은 아닌 것으로 분석되었다. 즉, 업종 별로 인스타그램과의 영향 정도는 다르다.

연관이 있는 업종

외식업

: 커피/음료, 한식, 제과점, 양식, 패스트푸드

소매업

: 패션용품, 의류점, 화초(꽃집)

서비스업

: 네일

WHY 🔍

커피/음료

- ✓ "Instagram-able"한 인테리어와 분위기가 형성
=> 인스타그램에 사진을 업로드하기 좋음

한식

- ✓ 망원시장에 많은 한식 업종 분포
- ✓ 망원시장: TV, 뉴스 등 다양한 매체에서 언급되며
인지도가 높아짐

② 분석 대상에 포함된 모든 업종이 인스타그램과 영향 관계에 있는가?

분석 대상의 모든 업종들이 인스타그램과 영향 관계에 있는 것은 아닌 것으로 분석되었다. 즉, 업종 별로 인스타그램과의 영향 정도는 다르다.

연관이 있는 업종

외식업

: 커피/음료, 한식, 제과점, 양식, 패스트푸드

소매업

: 패션용품, 의류점, 화초(꽃집)

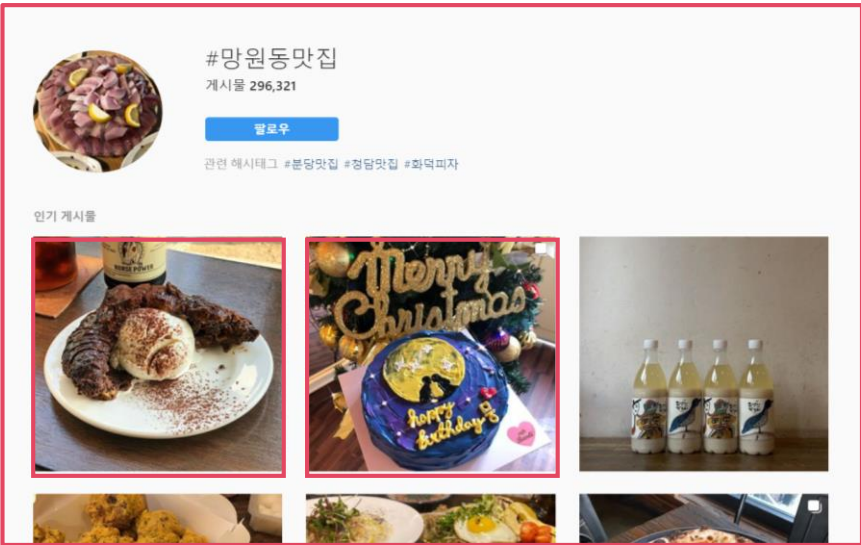
서비스업

: 네일

WHY 🔍

제과점

- ✓ 시각적으로 화려한 케이크, 디저트류가 많음
- ✓ 점포 수는 적지만(18개) 반응도가 뛰어남
- ✓ 케이크: 인스타그램을 통한 예약 주문 활성화



② 분석 대상에 포함된 모든 업종이 인스타그램과 영향 관계에 있는가?

분석 대상의 모든 업종들이 인스타그램과 영향 관계에 있는 것은 아닌 것으로 분석되었다. 즉, 업종 별로 인스타그램과의 영향 정도는 다르다.

연관이 있는 업종

외식업

: 커피/음료, 한식, 제과점, 양식, 패스트푸드

소매업

: 패션용품, 의류점, 화초(꽃집)

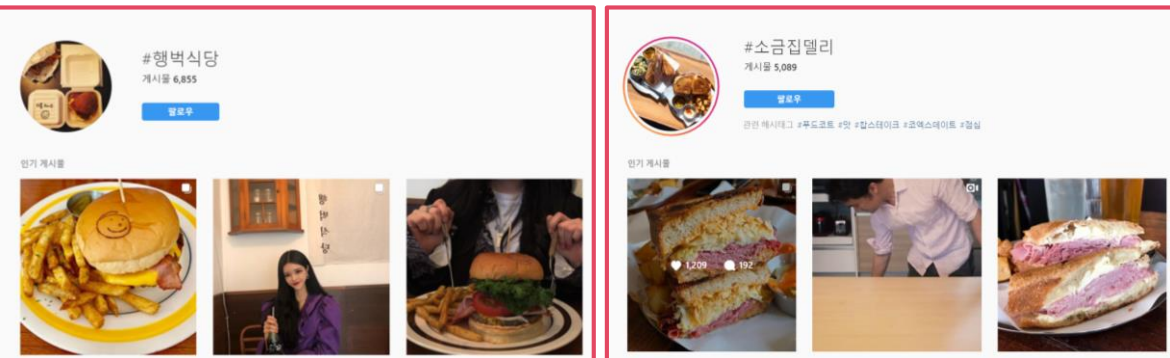
서비스업

: 네일

WHY 🔍

패스트푸드

- ✓ 망원에 프랜차이즈 패스트푸드점(ex. 맥도날드, 롯데리아)는 2개 밖에 없음
- ✓ 다른 곳에서 찾아보기 힘든 소규모 수제 패스트푸드점이 많음. (ex. 행복식당, 소금집델리)
→ 음식 비주얼+차별화 되는 상품+인플루언서들의 언급



② 분석 대상에 포함된 모든 업종이 인스타그램과 영향 관계에 있는가?

분석 대상의 모든 업종들이 인스타그램과 영향 관계에 있는 것은 아닌 것으로 분석되었다. 즉, 업종 별로 인스타그램과의 영향 정도는 다르다.

연관이 있는 업종

외식업

: 커피/음료, 한식, 제과점, 양식, 패스트푸드

소매업

: 패션용품, 의류점, 화초(꽃집)

서비스업

: 네일

WHY 🔍

패션용품

✓ 소규모 공방을 운영하는 문화예술인이 많았던 홍대상권에서 확장됨

=> 젠트리피케이션으로
홍대-합정-망원으로 이동한
소규모 공방들

=> 접근성에 크게 구애 받지 않는
업종



② 분석 대상에 포함된 모든 업종이 인스타그램과 영향 관계에 있는가?

분석 대상의 모든 업종들이 인스타그램과 영향 관계에 있는 것은 아닌 것으로 분석되었다. 즉, 업종 별로 인스타그램과의 영향 정도는 다르다.

WHY 🔍

- ✓ 마포구에 비해 호프점 비율이 크게 적다. 마포구 전체 모든 업종 점포 수 대비 호프 점포 수 비율은 8.96%
망원동 모든 업종 점포 수 대비 호프 점포 수 비율은 1.43%이다
=> 망원동은 마포구 평균에 비해 호프 점포수가 6~7배 가량 적다는 것을 알 수 있음
- ✓ 이는 망원에서 술집이 유의미한 수익을 올리지 못하고 있다는 것을 반증. 타지역에 비해 특색이 없기 때문일 것으로 예상함

연관이 없는 업종

외식업

: 중식, 일식, 호프/바

소매업

: -

서비스업

: 스포츠클럽, 미용실

③ 연도 별로 인스타의
영향력이 달라지는가?

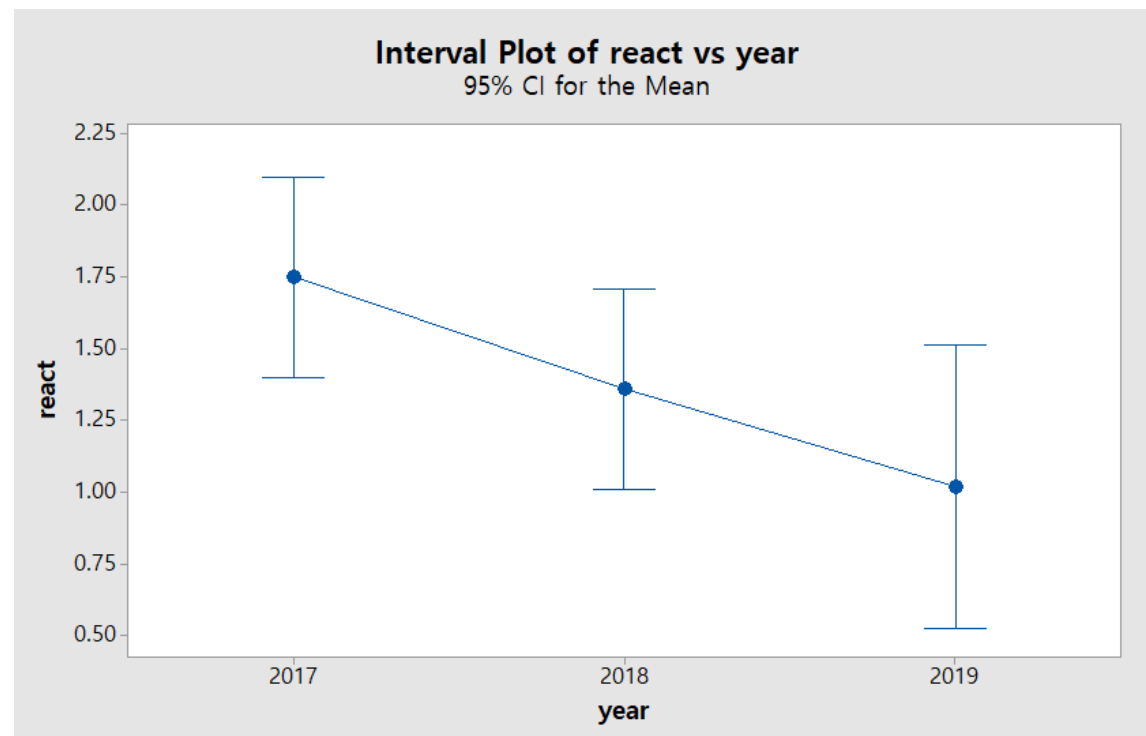
ANOVA 분산분석

ANOVA: react versus year

Factor	Type	Levels	Values
year	fixed	3	2017, 2018, 2019

Analysis of Variance for react

Source	DF	SS	MS	F	P
year	2	21.443	10.722	3.04	0.049
Error	277	976.649	3.526		
Total	279	998.092			

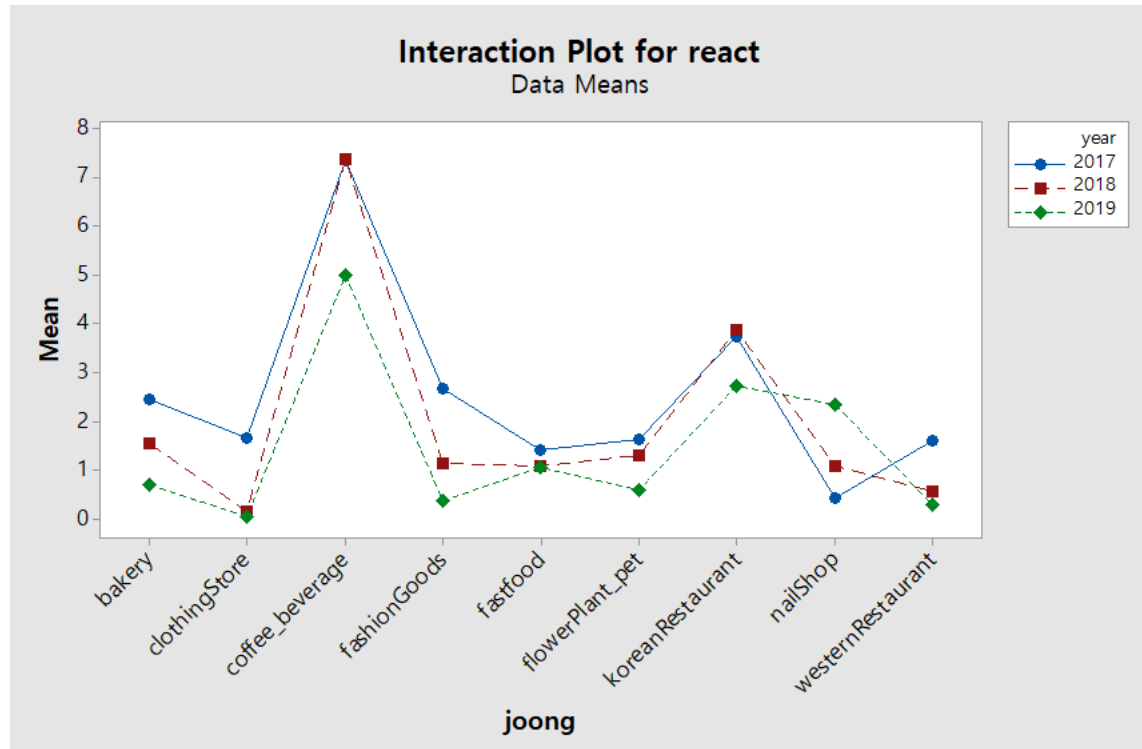


연도 변수의 수준(2017, 2018, 2019) 차이에 따른 반응도의 차이가 있다는 결과 도출

Interaction Plot (반응도X업종*연도)

종속변수: 반응도, 독립변수: 연도(2017, 2018, 2019)와, 반응도와 유의한 업종들의 교호작용을 나타내기 위한 분석 실시

최종 결과



대부분의 업종에서, 2017년에서 2019
년으로 갈 수록 반응도 값이 하락하고 있음

다만 네일샵의 경우 2019년으로 갈 수록
반응도 값이 상승하고 있음

③ 연도 별로 인스타의 영향력이 달라지는가?

연도 별로 인스타의 영향력은 달라지는 것을 확인할 수 있었다. 즉, 2017년에 비해 2019년으로 갈 수록, 업종과 반응도와의 영향관계가 줄어들고 있다고 분석되었다.

WHY 🔍

Geographic

홍대와의 인접성으로
많은 유동인구 유입

Economic

홍대에 비해 상대적으로
저렴한 임대료

Socio-Culture

프랜차이즈에 대한 피로도
과시를 즐기는 사회

③ 연도 별로 인스타의 영향력이 달라지는가?

연도 별로 인스타의 영향력은 달라지는 것을 확인할 수 있었다. 즉, 2017년에 비해 2019년으로 갈 수록, 업종과 반응도와의 영향관계가 줄어들고 있다고 분석되었다.

WHY 🔍

한계 1

모방 가능성이
높은 산업

한계 2

인스타그램으로 유입된
소비자 특성 => "일회적"

한계 3

높아지는 인기와 더불어
임대료 상승

③ 연도 별로 인스타의 영향력이 달라지는가?

연도 별로 인스타의 영향력은 달라지는 것을 확인할 수 있었다. 즉, 2017년에 비해 2019년으로 갈 수록, 업종과 반응도와의 영향관계가 줄어들고 있다고 분석되었다.

SO 

IMPLICATION

✓ 해당 상권은 트렌드가 굉장히 빠르다
: 1년 주기로 업종의 반응도가 떨어지고 있다

✓ 따라서 2020년에도 인스타그램을 활용한
홍보가 유효할지에 대해서 장담할 수 없다

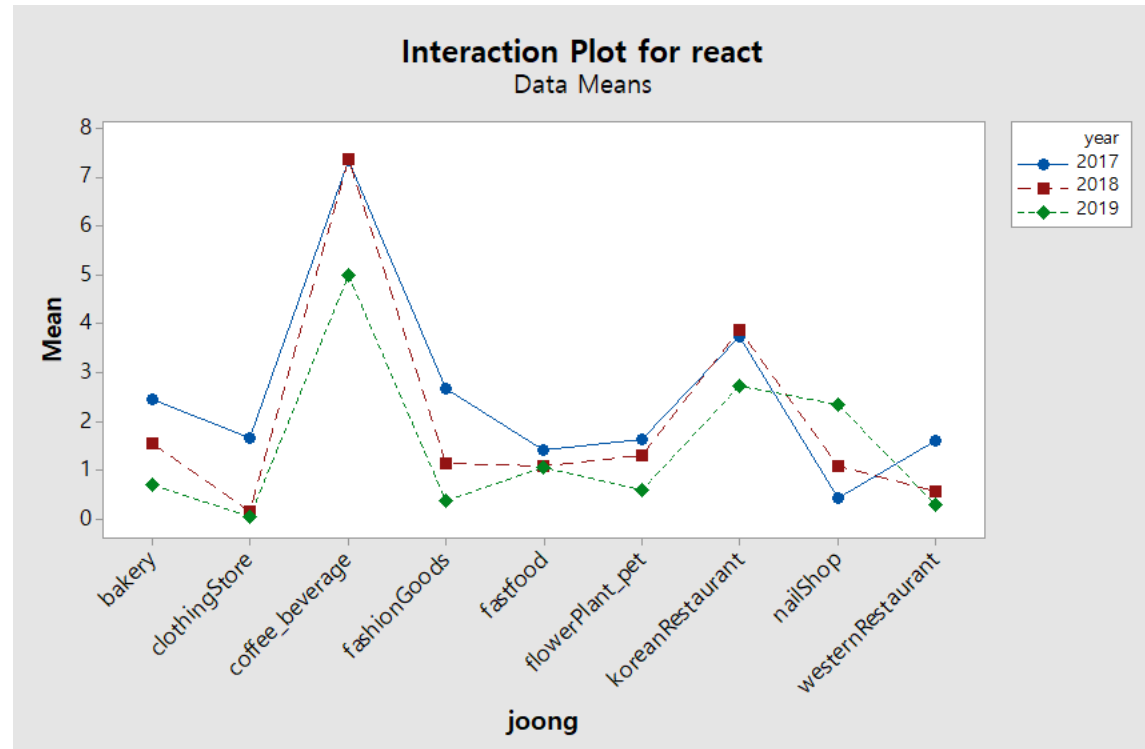
③ 연도 별로 인스타의 영향력이 달라지는가?

연도 별로 인스타의 영향력은 달라지는 것을 확인할 수 있었다. 즉, 2017년에 비해 2019년으로 갈 수록, 업종과 반응도와의 영향관계가 줄어들고 있다고 분석되었다.

Nail 🔍

IMPLICATION

- ✓ 산업안전보건연구원 통계에 따르면 국내 네일샵은 2015년에 6000개 수준이었으나, 2018년 1만3258개로 3년 만에 약 120%를 웃도는 급성장을 하고 있다.
- ✓ 네일 업종이 해마다 급격한 성장을 이루고 있음. 네일은 연예인, 셀럽에 의해 크게 영향을 받는 업종이며, 취향을 많이 타는 업종임
- ✓ 다양한 디자인을 볼 수 있는, 셀럽들의 사진이 많이 올라오는 인스타그램 플랫폼과 연관 있을 수 밖에 없음



④ 분기 별로 세분화하여 볼 때,
인스타그램의 영향도는
특정 분기 별로 달라지는가?

ANOVA 분산분석

One-way ANOVA: react versus bungi

Method

Null hypothesis All means are equal
 Alternative hypothesis At least one mean is different
 Significance level $\alpha = 0.05$

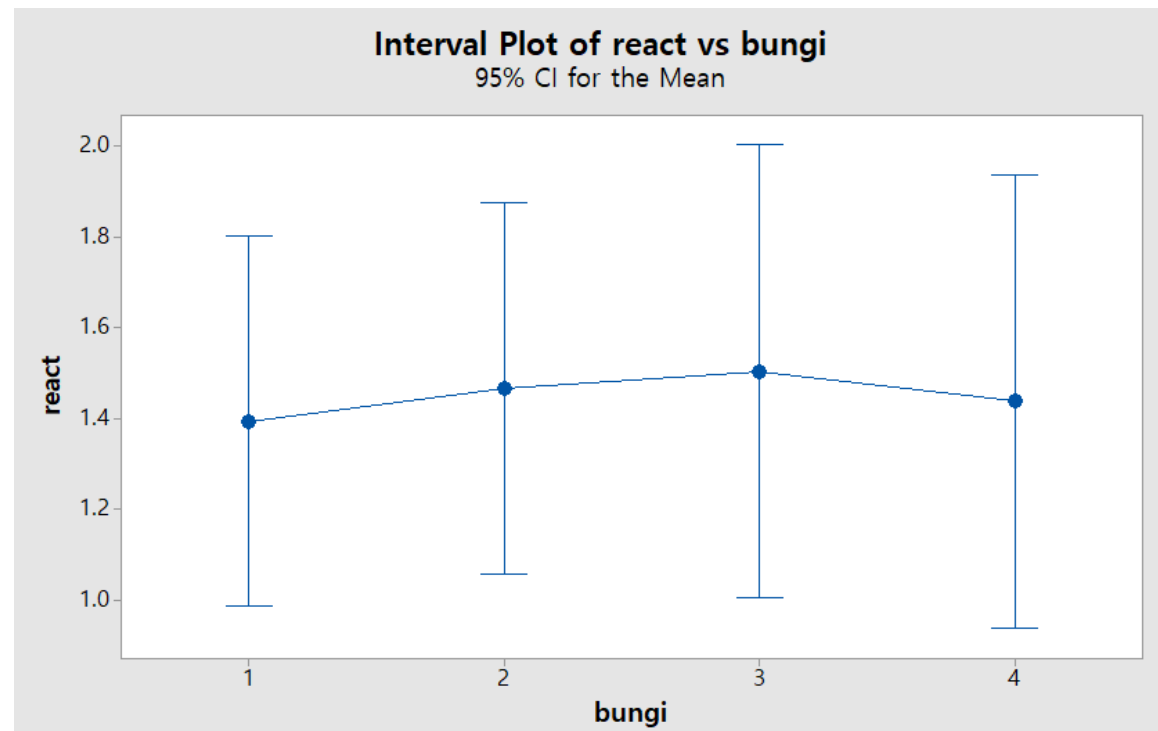
Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
bungi	4	1, 2, 3, 4

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
bungi	3	0.459	0.1531	0.04	0.988
Error	276	997.633	3.6146		
Total	279	998.092			



분기 변수의 수준(1분기, 2분기, 3분기, 4분기) 차이에 따른 반응도의 차이가 없다는 결과

회귀분석

종속변수: 반응도, 독립변수: 각 분기로 설정 후 회귀분석 시행

```
> summary(gggg)
```

```
Call:
lm(formula = react ~ bungi_1 + bungi_2 + bungi_3 + bungi_4, data = our)
```

Residuals:

Min	1Q	Median	3Q	Max
-1.5039	-1.2505	-0.6754	0.4703	8.4938

Coefficients: (1 not defined because of singularities)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1.43750	0.25406	5.658	3.83e-08 ***
bungi_1	-0.04417	0.32799	-0.135	0.893
bungi_2	0.02869	0.32799	0.087	0.930
bungi_3	0.06643	0.35930	0.185	0.853
bungi_4	NA	NA	NA	NA



유의하지 않다는 결과가 나왔다

=> 분기별로 반응도의 영향이 다르다는 가설은 틀림

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

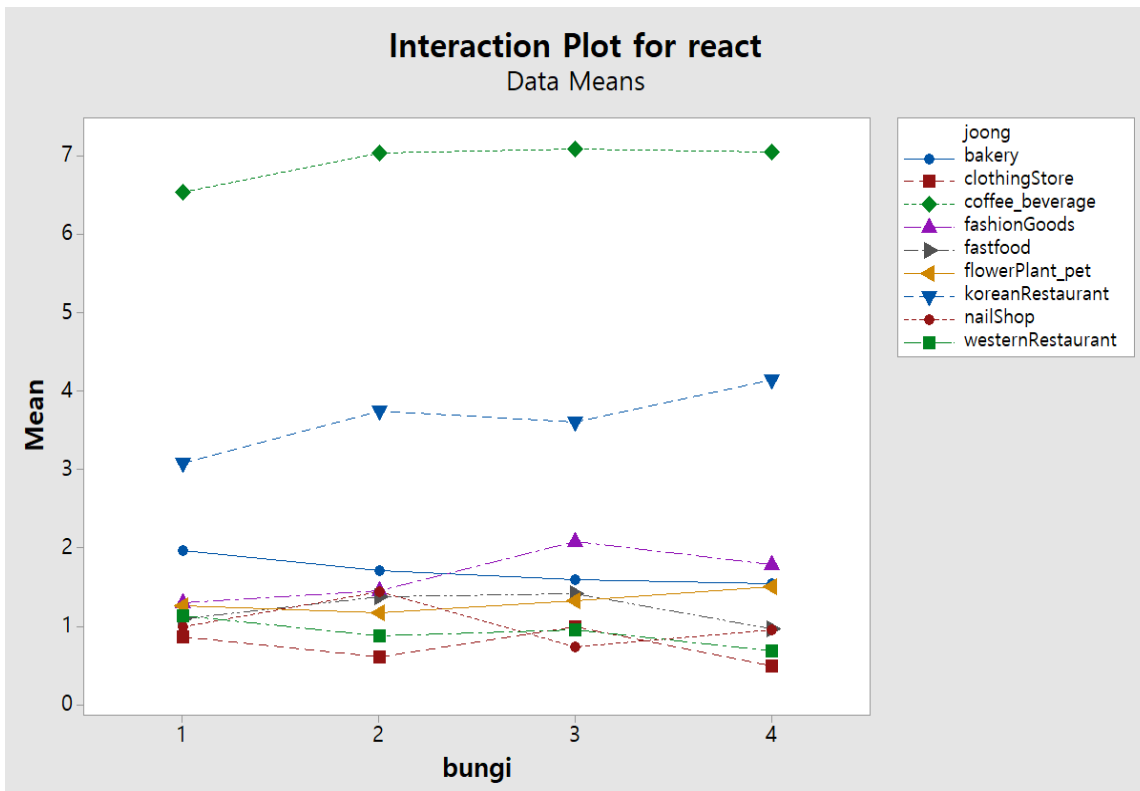
Residual standard error: 1.901 on 276 degrees of freedom

Multiple R-squared: 0.0004601, Adjusted R-squared: -0.0104

F-statistic: 0.04235 on 3 and 276 DF, p-value: 0.9884

Interaction Plot (반응도X분기*업종)

종속변수: 반응도, 독립변수: 분기(1분기, 2분기, 3분기, 4분기)와, 반응도와 유의한 업종들의 교호작용을 나타내기 위한 분석 실시



계절에 따라서 (분기에 따라서)
반응도 지표는 달라지지 않았음

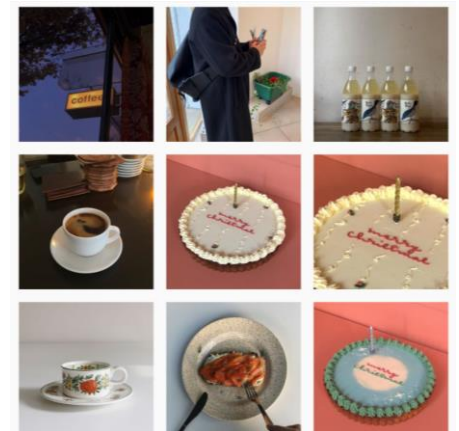
④ 인스타그램의 영향도는 특정 분기 별로 달라지는가

분기 별로 인스타그램의 영향도가 달라질 것이라는 가설은 기각되었다. 즉, 계절 요인은 특정 업종에 대한 인스타그램의 반응에 관하여 영향을 미치지 못한 것으로 분석되었다.

WHY 🔍

인스타그램으로 유입된 소비자들은 **계절 요인과는 무관한 특정 목적을 가지고 방문하기 때문임**

- ✓ 계절 요인에 따라 특정 업종에 대한 인스타그램 반응이 달라지는지 확인하고자 했음 → 연관 X
- ✓ #망원동으로 검색했을 때 인기 게시물: 예쁜 사진 + 모두 1000명 이상의 팔로워를 가진 인플루언서의 게시물
- ✓ 표본 업종에서 계절 요인을 강하게 받는 업종도 없었음



3

분류 분석 모델

분류 분석 모델

앞서 회귀 분석 실시 결과 **반응도와 유의하다는 결과가 나온 업종들**을 반응도에 영향을 많이 끼치는 업종들로 **군집화**

커피/음료, 한식, 제과점, 패션용품, 중식, 미용실, 호프/바, 일식, 의류점

-> 인스타그램 영향 많이 받는 업종 -> **Labeling 1**

그렇지 않은 업종들을 **군집화** 하여 0으로 Labeling

스포츠클럽, 네일샵, 화초/애완, 양식, 패스트푸드

-> 인스타그램 영향을 받지 않는 업종 -> **Labeling 0**

분류 분석 모델

사용 변수, 프랜차이즈 점포 수, 일반 점포 수, 개업률, 폐업률, 10년 평균 영업기간 등 업종을 나타내는 지표

	joong	FranStores	GeneralStores	OpeningRatio	ClosingRatio	last10AveragePeriod
0	0	0.0	0.042857	0.437063	0.000000	0.300000
1	0	0.0	0.014286	0.000000	0.874126	0.566667
2	1	0.0	0.350000	0.069930	0.000000	0.966667
3	1	0.1	0.421429	0.055944	0.000000	0.900000
4	0	0.0	0.035714	0.000000	0.000000	0.566667

분류 분석 모델

훈련집합 평가 집합 8대 2로 나누기

```
# 훈련집합, 평가집합 나누기
from sklearn.model_selection import train_test_split
train_x, test_x, train_y, test_y = train_test_split(ourdata.drop('joong', axis=1), ourdata['joong'], test_size=0.2, random_state=42)
```

```
train_x.shape
```

```
(224, 5)
```

```
test_x.shape
```

```
(56, 5)
```

분류 분석 모델

모델 평균 정확도 0.91

```
# RandomForestClassifier
from sklearn.ensemble import RandomForestClassifier

model = RandomForestClassifier(random_state=13)
model.fit(train_x, train_y)
```

```
# predict 메서드를 호출하여 test_x의 값으로 모델을 테스트한 다음,
# score 메서드를 사용하여 모델의 평균 정확도를 확인합니다.
predicted = model.predict(test_x)
model.score(test_x, test_y)
```

0.9107142857142857

혼동행렬 결과 값

오류 값이 5개밖에
발견되지 않음

```
# 혼동 행렬 생성 - 혼동 행렬은 각 답변이 정확하거나 잘못 분류된 횟수를 정량화합니다.
# 특히 거짓 긍정, 거짓 부정, 참 긍정 및 참 부정의 수를 정량화합니다.
from sklearn.metrics import confusion_matrix
confusion_matrix(test_y, predicted)
```

```
array([[18,  3],
       [ 2, 33]], dtype=int64)
```

분류 분석 모델

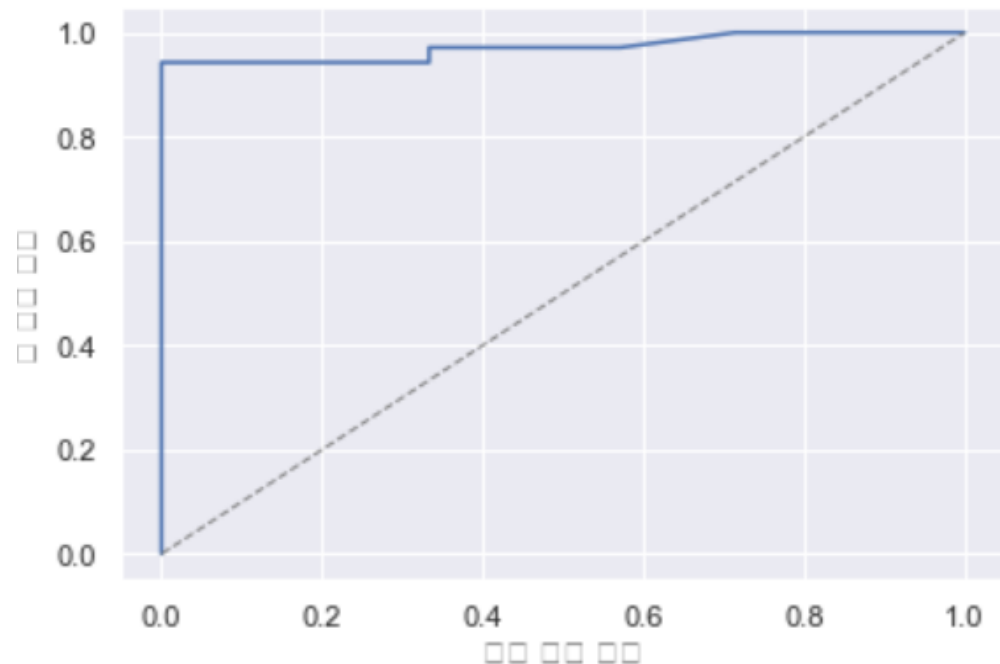
```
roc_auc_score(test_y, probabilities[:, 1])
```

0.972108843537415

ROC AUC가 0.97정도로 매우 이상적인 값이 나옴

```
# ROC 곡선
from sklearn.metrics import roc_curve

fpr, tpr, _ = roc_curve(test_y, probabilities[:, 1])
plt.plot(fpr, tpr)
plt.plot([0, 1], [0, 1], color='grey', lw=1, linestyle='--')
plt.xlabel('거짓 긍정 비율')
plt.ylabel('참 긍정 비율')
```



분류 분석 모델

반응도를 구하지 못했던 업종들의 19년도 2분기 데이터를 입력 변수로 넣어 인스타그램의 반응도를 예측해 봄

중분류	franNums	GeneNum	openRatio	closeRatio	10yearsAver
일반교습학원	0.045455	0.148148	0.000000	0.000000	0.428571
외국어학원	0.045455	0.111111	0.000000	0.000000	0.452381
예체능학원	0.045455	0.259259	0.000000	0.000000	0.452381
치과의원	0	0.166667	0.000000	0.000000	0.785714
한의원	0	0.240741	0.000000	0.000000	0.357143
일반의원	0	0.388889	0.210000	0.000000	0.071429
가전제품수리점	0	0.055556	0.000000	0.380000	0.714286
부동산중개업	0	0.925926	0.000000	0.000000	0.428571
숙박업	0	0.000000	0.000000	0.000000	1
노래방	0	0.277778	0.000000	0.000000	0.190476
PC방	0.045455	0.018519	1.000000	0.000000	0.238095
당구장	0	0.129630	0.500000	0.000000	0.166667
세탁소(가정용)	0.090909	0.203704	0.000000	0.000000	0.690476
자동차수리점	0	0.185185	0.000000	0.000000	0.761905
피부관리샵	0	0.296296	0.000000	0.510000	0
슈퍼마켓	0	0.666667	0.640000	0.000000	0.071429
편의점	1	0.018519	0.100000	1.000000	0.190476

	franNums	GeneNums	openRatio	closeRatio	10yearsAver
0	0.045455	0.148148	0.0	0.0	0.428571
1	0.045455	0.111111	0.0	0.0	0.452381
2	0.045455	0.259259	0.0	0.0	0.452381
3	0.000000	0.166667	0.0	0.0	0.785714
4	0.000000	0.240741	0.0	0.0	0.357143

```
predicted2 = model.predict(test_X2)
predicted2
```

```
array([0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1,
       1, 1, 1, 0, 0, 1, 1, 1, 1], dtype=int64)
```

분류 분석 모델

반응도를 구하지 못했던 업종들의 19년도 2분기 데이터를 입력 변수로 넣어 인스타그램의 반응도를 예측해 봄

	franNums	GeneNums	openRatio	closeRaito	10yearsAver
0	0.045455	0.148148	0.0	0.0	0.428571
1	0.045455	0.111111	0.0	0.0	0.452381
2	0.045455	0.259259	0.0	0.0	0.452381
3	0.000000	0.166667	0.0	0.0	0.785714
4	0.000000	0.240741	0.0	0.0	0.357143

```
predicted2 = model.predict(test_X2)
predicted2
```

```
array([0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1,
       1, 1, 1, 0, 0, 1, 1, 1, 1], dtype=int64)
```

pc방 0	섬유제품 1	자동차수리/세차 1
가구/가전 0	세탁소(가정) 0	주방/가정용품 1
가전제품수리 1	숙박업 0	치과의원 1
건강보조식품 1	슈퍼마켓 1	치킨전문점 1
노래방 0	식료품 1	컴퓨터/주변기기 1
당구장 1	예체능학원 1	통신판매업 0
부동산중개업 1	오락/운동 1	편의점 0
분식전문점 1	외국어학원 0	피부관리실 1
서적/문구 0	의약/의료용품 0	한의원 1
일반교습학원 1	일반의원 1	핸드폰 1
		화장품 1

0인 업종들 -> 인스타그램의 영향력을 **덜** 받는 업종

1인 업종들 -> 인스타그램의 영향력을 **잘** 받는 업종

1인 업종들은 인스타그램으로 홍보를 해도 될 것임!!!

Thank you 🥰