



에이콘아카데미_0127_파이널_빅데이터분석과정

How

About

This?

HAT_TAG

양질의 콘텐츠 구성을 위한

적절한 해시태그 제안

01 서론

1. 팀 소개
2. 주제 소개
 - 2-1 주제
 - 2-2 왜 인스타그램?
 - 2-3 Target 설정
3. 진행 과정
 - 3-1 분석과정
(사용 라이브러리, 분석방법)
 - 3-2 데이터 수집
 - 3-3 결과물 (Model, DF)

02 본문

1. 트렌드 분석
 - 1-1 좋아요 상위 300개
2. 음식별 분석
 - 2-1 치킨 해시태그
3. 게시물 유형 분석
 - 3-1 피드 올리는 방법
4. 통계적 분석
 - 4-1 필수태그 / 추가태그
5. 모델 기반 추천 태그

03 결론

1. 결론
 - 1-1 구현
 - 1-2 종합
2. 아쉬운 점

<https://finalproject/finale/서론>

Instagram



홈



검색



탐색 탭



릴스



메시지



알림



만들기



프로필



더 보기

검색

검색

최근 검색 항목

모두 지우기

#

#1. 팀 소개

X

#

#2. 주제 소개

X

2-1 주제

2-2 왜 인스타그램?

2-3 Target 설정

#

#3. 진행 과정

X

3-1 분석과정

3-2 데이터 수집

3-3 결과물

(Model, DataFrame)

인기 게시물





팀원 소개

팀장

전대한

이미지 데이터 크롤링
이미지 분류 모델 설계
데이터 전처리
ppt 제작
발표

팀원

박주현

이미지 데이터 크롤링
이미지 분류 모델 설계
시각화
ppt 제작
보고서 작성

팀원

임성길

이미지 데이터 크롤링
이미지 분류 모델 설계
시각화
데이터 분석
Ppt개요 구성

팀원

남수민

텍스트 데이터 크롤링
데이터 전처리
시각화
상관 분석
본론 분석

팀원

이소민

텍스트 데이터 크롤링
데이터 전처리
본론 분석
통계 분석

팀원

박명훈

텍스트 데이터 크롤링
데이터 전처리
시각화
최종 분석



사용 라이브러리



데이터 수집

이미지 데이터 크롤링
텍스트 데이터 크롤링



모델링

이미지 데이터 전처리
모델 설계



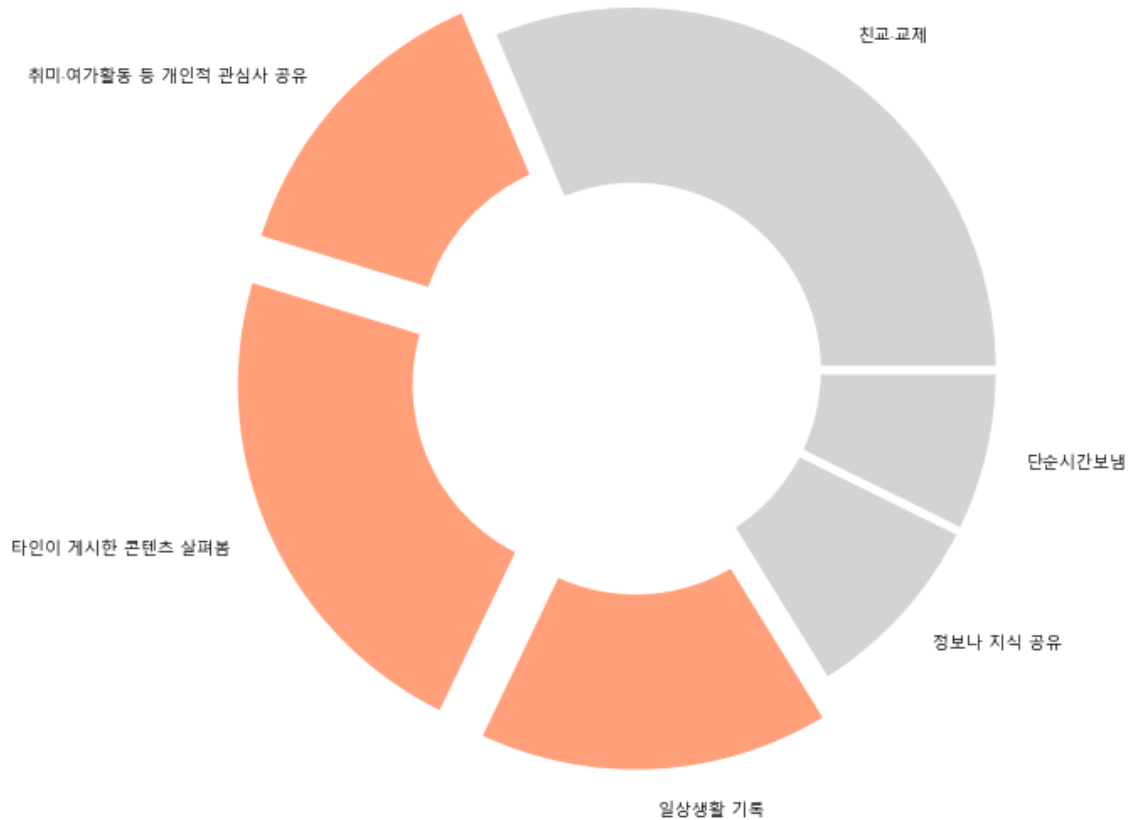
데이터 분석

텍스트 데이터 전처리
시각화
통계적 분석



왜 인스타그램?

<SNS 이용 이유>



인스타, SNS·커뮤니티 앱 월간 이용자 국내 첫 1위...네이버 밴드 제쳐

입력 2022.08.04 06:00 수정 2022.08.04 06:00

가가

출처 : 한경미디어_22년8월4일

인스타그램, 밴드 앞지를까?...12월 설치자 약 28만차

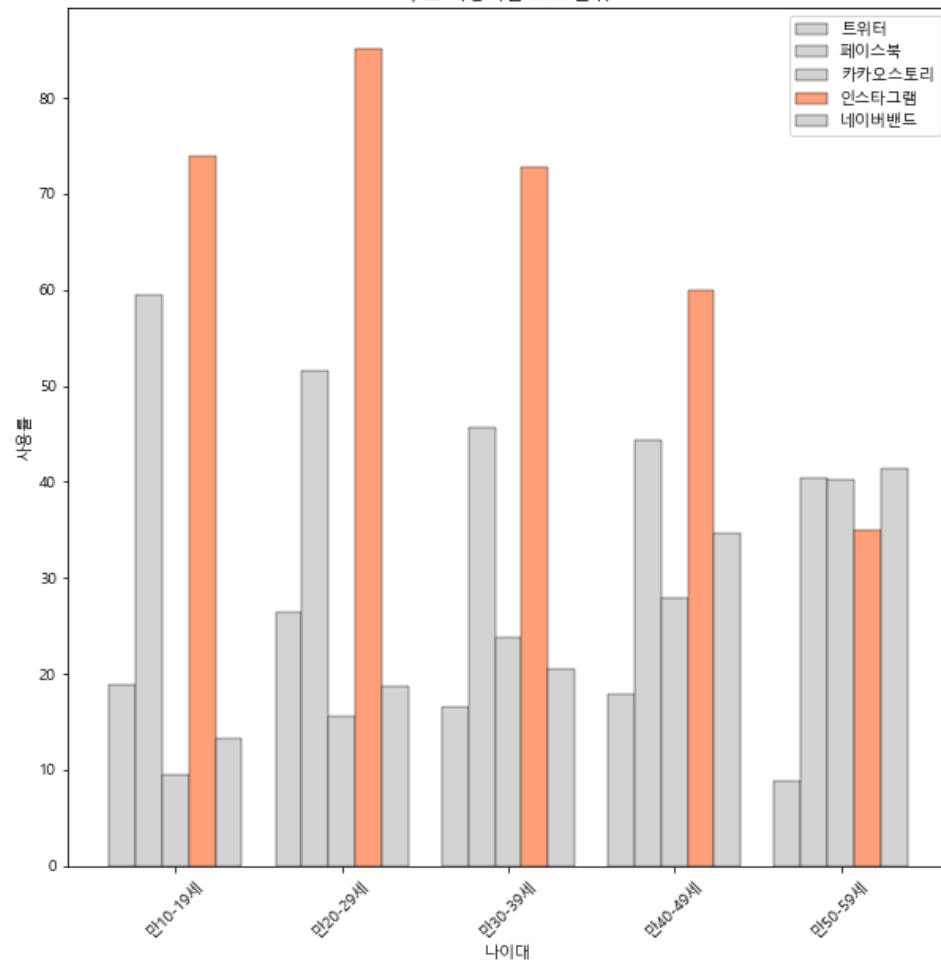
입력 : 2023-01-12 11:04:26 | 수정 : 2023-01-12 11:04:24

출처 : 세계일보_23년1월12일

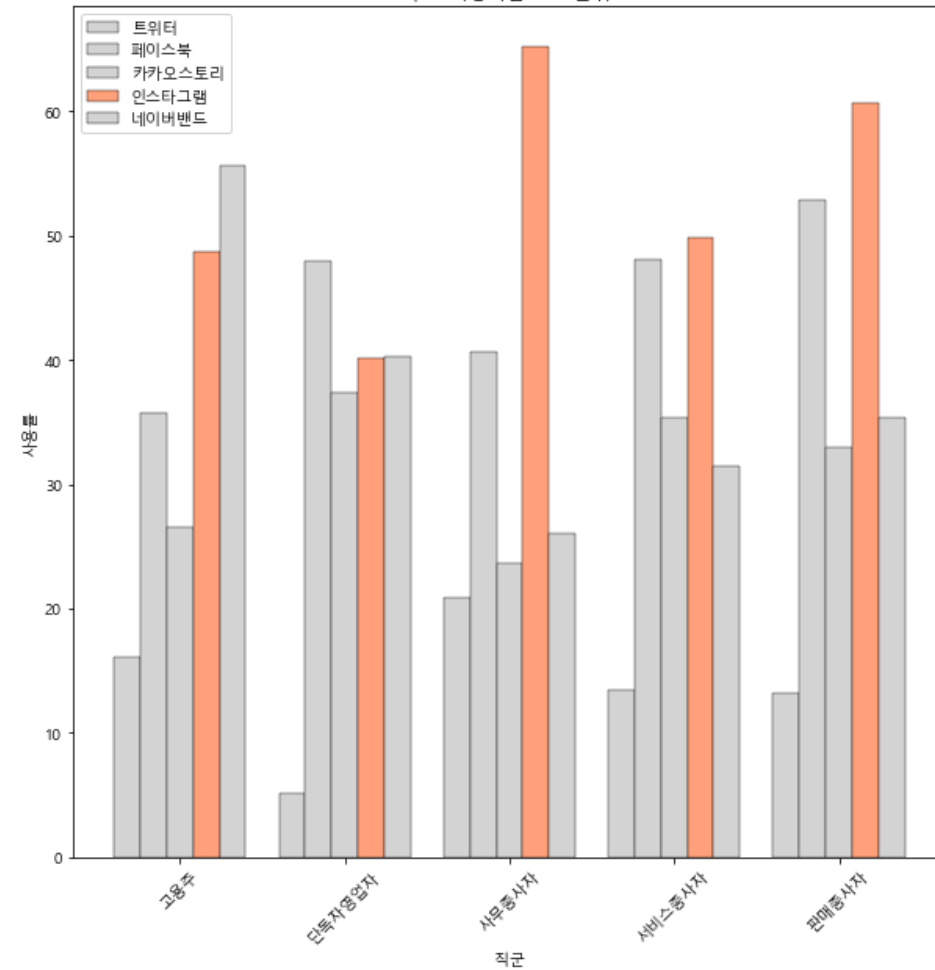


왜 인스타그램?

주로 사용하는 SNS 순위



주로 사용하는 SNS 순위





왜 외식 메뉴?

광고 집행 주요 업종 비교

성공적인
SNS 광고를 위한
마케팅 가이드

- 공통적으로 다음 업종들이 SNS 광고를 선호 - **식음료/식품, 엔터테인먼트, 패션/의류, 화장품/보건용품**
- 이용 유저와 서비스 특성에 따라 업종 순위의 차이가 존재, SNS별 업종 순위를 참고하여 미디어 전략을 구성할 것을 추천



페이스북

1위

패션/의류

2위

식음료/식품

3위

엔터테인먼트

4위

제약/의료

5위

화장품/보건용품



인스타그램

식음료/식품

화장품/보건용품

패션/의류

엔터테인먼트

전기전자



밴드

유통

교육/복지후생

서비스

금융

전기전자



트위터

컴퓨터/정보통신

식음료/식품

엔터테인먼트

그룹/기업광고

서비스



틱톡

식음료/식품

엔터테인먼트

패션/의류

컴퓨터/정보통신

화장품/보건용품



Target 설정

Micro-Influencer란?

Influencer : 영향력을 행사하는 사람, 감화시키는 사람

1. 소셜 미디어에서 특정 분야의 콘텐츠를 많이 만들고 다른 사람들에게 영향력을 발휘하는 사람
2. 인스타그램 내에서 약 500~10만명 이내의 팔로워를 보유한 사람들을 의미
3. 인플루언서와 팔로워의 상호작용성 정도에 따라 광고효과의 차이를 보임

-> 비교적 팔로워들과의 즉각적인 소통과 긴밀한 관계를 유지할 수 있는 “Micro-Influencer”를 통한 광고 효과가 검증되고, “#광고” 라는 태그에 대해서 스폰서십 공시의 부정적인 효과가 그 상호작용성에 의해 상쇄될 수 있다.

#인스타그램 #상호작용 #제품소개 #광고



분석 과정

텍스트 데이터



이미지 데이터

이미지 분류를 통한
태그 추천



데이터 수집(이미지)

사용 모듈

BeautifulSoup



크롤링

kaggle

Google

NAVER

```
import  
CV, OS, SHUTIL,
```

최종 폴더



Chicken / 3580장



Pizza / 4813장



Tteokbokki / 3695장



Pasta / 3562장



Sushi / 3650장



Burger / 4024장



Porkcutlet / 4756장



Train / 28,080장



이미지 분류 모델 설계

최종 폴더



Chicken / 3580장



Pizza / 4813장



Tteokbokki / 3695장



Pasta / 3562장



Sushi / 3650장



Burger / 4024장



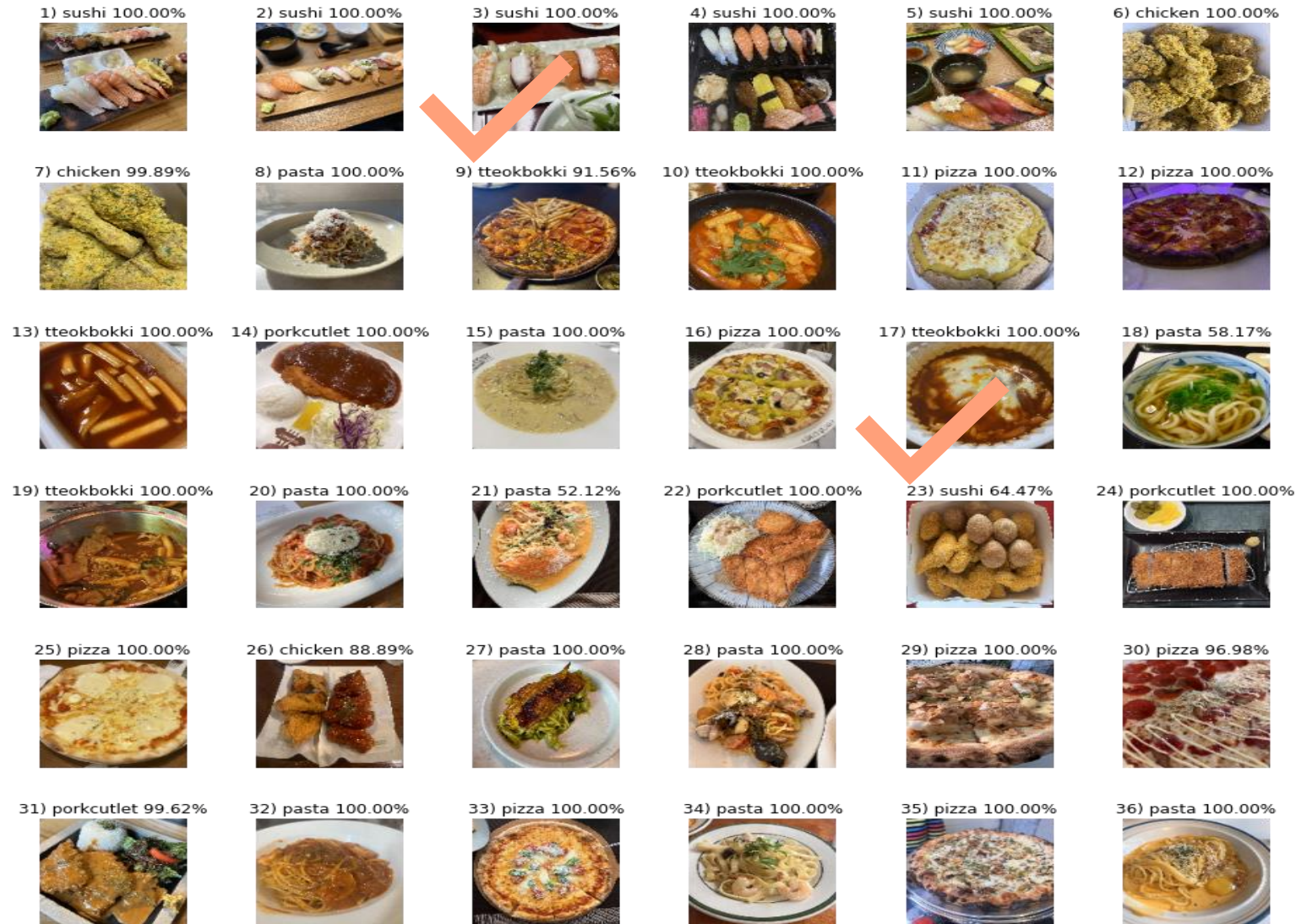
Porkcutlet / 4756장

모델설계

모델설계	훈련점수	테스트점수	최종모델
Mobilenet + 은닉층 1개(units=64 / lr=0.01)	X
Mobilenet + 은닉층 2개(64,128 / lr=0.01)	X
Resnet50 + 은닉층 2개(64,128,256 / lr=0.01)	X
Mobilenet + 은닉층 3개(64,128,256 / lr=0.0001)	99.72	96.72	O
Mobilenet + 은닉층 4개(lr=0.01)	X
Mobilenet + 은닉층 2개(64,128 / lr=0.0001)	X

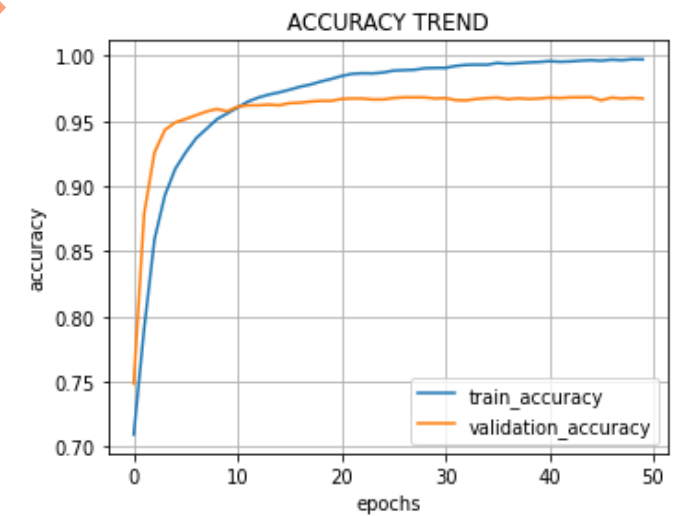


이미지 분류 모델 설계



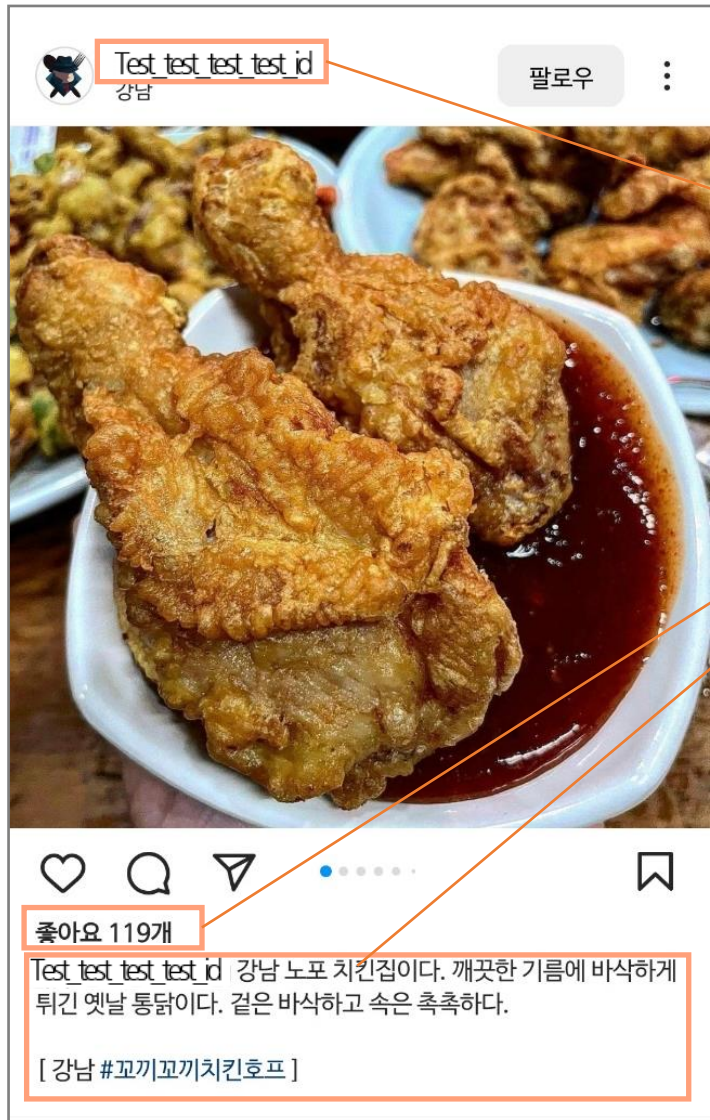
최종 모델

1. MobileNet
2. Lr(학습률) = 0.00001
3. 은닉층 = 3개
4. Epochs = 50
5. Batch_size = 128





데이터 수집(텍스트)



1차 크롤링

ID (인스타그램 계정 아이디)

Contents (게시물 본문)

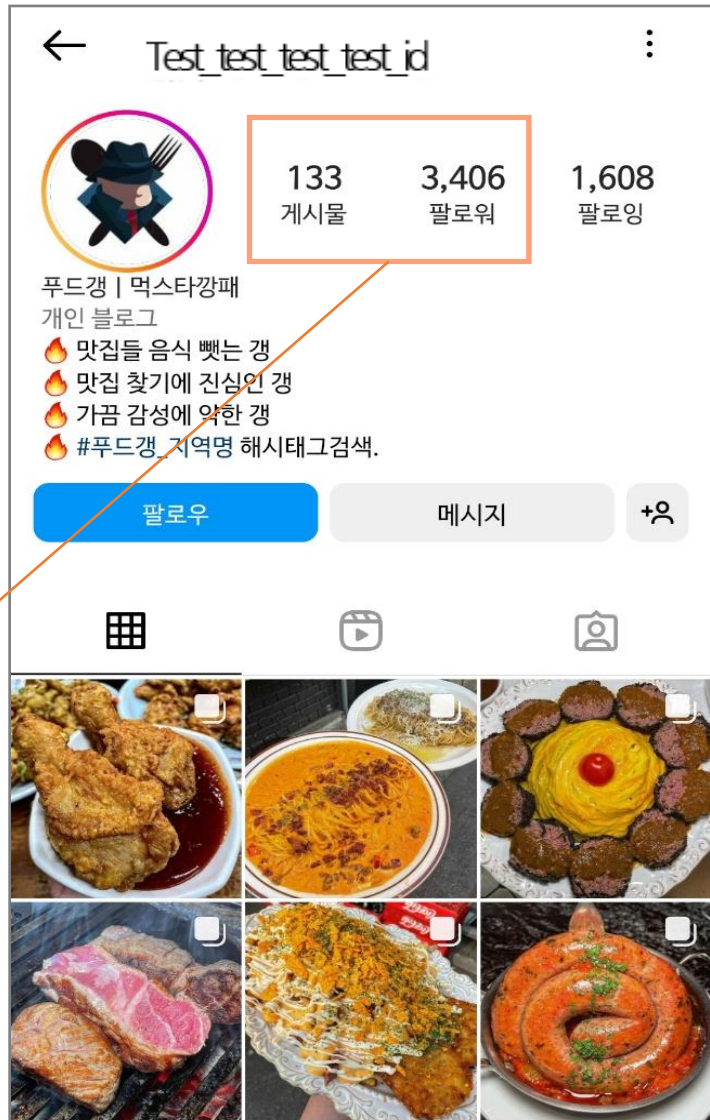
Tags(게시물 해시태그)

Likes(좋아요 수)

2차 크롤링

Post_num (게시물 수)

Followers (팔로워 수)





데이터 수집(텍스트)

전처리

변수 이름	전처리 방향	비고
Category(음식이름)	컬럼 추가	
ID(인스타그램 계정)	Category로 1차 크롤링	
Post_num(게시글수)	ID로 2차 크롤링	
Followers(팔로워수)	ID로 2차 크롤링	
Contents(게시글 본문)	Category로 1차 크롤링	정규표현식으로 특수문자 삭제
Tags(게시글 해시태그)	Category로 1차 크롤링	정규표현식으로 #을 기준으로 추출
Tags_sum(Tags의 수)	컬럼 추가	
Tags_count(빈도수 태그와 중복 수)	컬럼 추가	
Tags_rate(태그 중복 비율)	컬럼 추가	$(\text{Tags_count} / \text{Tags_sum}) * 100$
Likes(좋아요수)	Category로 1차 크롤링	'가장 먼저 좋아요를 눌러보세요' → 0 'oo님이 좋아합니다' → 1 '여러명이 좋아합니다' → 중간값으로 대체
Scaled_likes(Likes 정규화)	컬럼 추가	
Contents_len(Contents의 글자수)	컬럼 추가	



최종 데이터셋

Category	ID	Post_num	Followers	Contents	Tags	Tags_sum	Tags_count	Tags_rate	Likes	Scaled_likes	Contents_len
치킨	z.e.o.m	236.0	8603.0	동근해 대신 올리는 동근 해가 있습니다 우동사리에 치밥까지 치킨으로 줄코스 줄 기...	#동근이숯불두마리치 킨. #동근이숯불두마 리치킨. #치킨. #치킨 맛집. #배달치킨. #...	11	3	27.27	1040	0.175646	477
치킨	foodgangster_fg	128.0	3360.0	협찬 먹방이여 타피 오카 킬 쏘은 줄 알았 다 버블티같은 먹방 이다 먹주를 부르는...	#치킨메니아. #푸드경 배달. #배달맛집. #치 킨맛집. #먹방이맛집. #배달음식추천...	9	3	33.33	109	0.018409	262
치킨	bboong_bbang_2	741.0	6928.0	돼지 사료 알지 506 지코바 성장점빵빵스 가 스토레스 받으면 주기적 수월하는 ...		0	0	0.00	109	0.018409	511
치킨	dhriksptls	2178.0	27000.0	얼마나 맛있었으면 포장까지 ㅋㅋㅋ 별책 주 임		0	0	0.00	9	0.001520	26
치킨	a_a_b0s2om	606.0	1020.0	노랑통닭... 대구 맛집 먹방 instafood 요리스타그램 요리 맛스타...	#대구맛집. #먹방. #instafood. #요리스타 그램. #요리. #맛스타 그램. ...	28	8	28.57	19	0.003209	193

13,330개의 데이터
12개의 변수 및 파생변수

<https://finalproject/finale/본론>

Instagram



홈



검색



탐색 탭



릴스



메시지



알림



만들기



프로필



더 보기

검색

검색

최근 검색 항목

모두 지우기

#

#1. 트렌드 분석

X

1-1 좋아요 상위 300개

#

#2. 음식별 분석

X

2-1 치킨 해시태그

#

#3. 게시물 유형 분석

X

3-1 피드 올리는 방법

#

#4. 통계적 분석

X

4-1 필수태그 / 추가태그

#

#5. 모델 기반 추천 태그

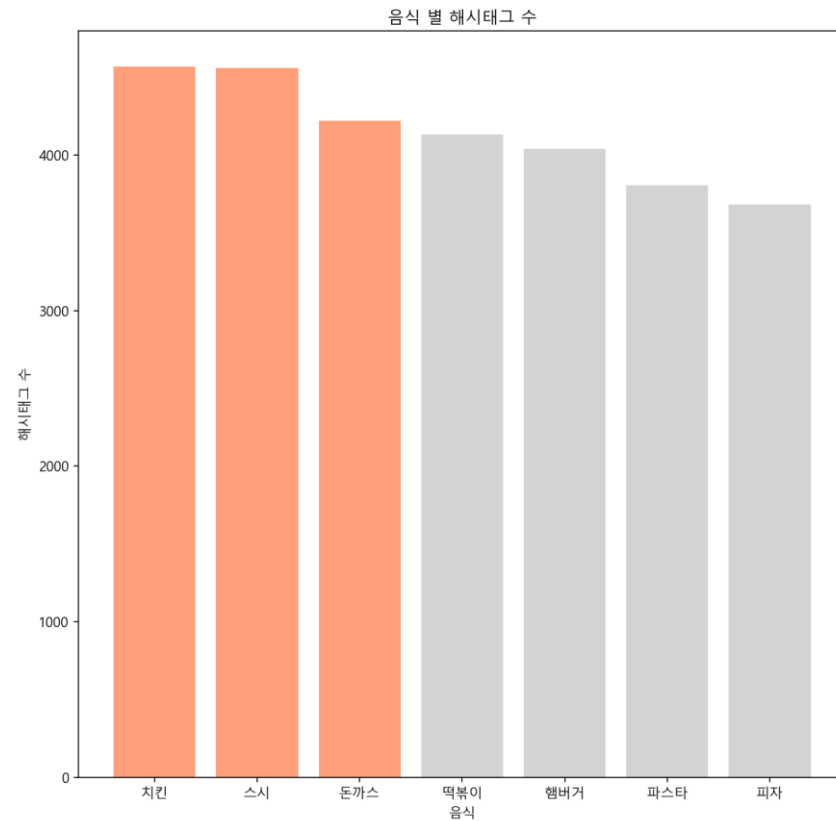
X

인기 게시물





2022 음식 트렌드



음식별 좋아요 수 상위 300개 데이터셋

#치킨 카테고리에서 사용된 해시태그의 개수는 총 **4568개**

#스시 카테고리에서 사용된 해시태그의 개수는 총 **4559개**

#돈까스 카테고리에서 사용된 해시태그의 개수는 총 **4219개**



트렌드 분석

#치킨

#저녁 #야식
#떡볶이 #맥주 #소주 #라면술과 함께 먹는 경우가 대다수
BHC, 교촌

#스시

#점심 #저녁
#사시미 #우동 #사케 #라면점심, 저녁메뉴로 소비
다른 메뉴와 함께 술안주로도 선호
오마카세 열풍을 통해 고급화를 추구

#돈까스

#점심 #저녁
#떡볶이 #우동 #짬면 #새우튀김일식과 경양식으로 나누어짐
주로 점심 혼밥 메뉴로 소비



2023 음식 트렌드 전망

치킨

떡볶이 혹은 라면(볶음면) 등과 함께 **다양한 조합**으로 소비될 것

스시

SNS 인증 특성을 고려한다면 이러한 **고급화 추세**가 더욱 가속화 될 것

돈까스

경양식은 추억의 맛, 일식은 정통성을 강조하며 **혼밥, 점심 메뉴**로 소비될 것

떡볶이

페어링 가능성이 높은 만큼 다양한 음식과의 **콜라보레이션**이 늘어날 것

햄버거

"수제"를 강조하고 있는 만큼 **수제햄버거**의 종류와 질 모두 증가할 것으로 예상

파스타

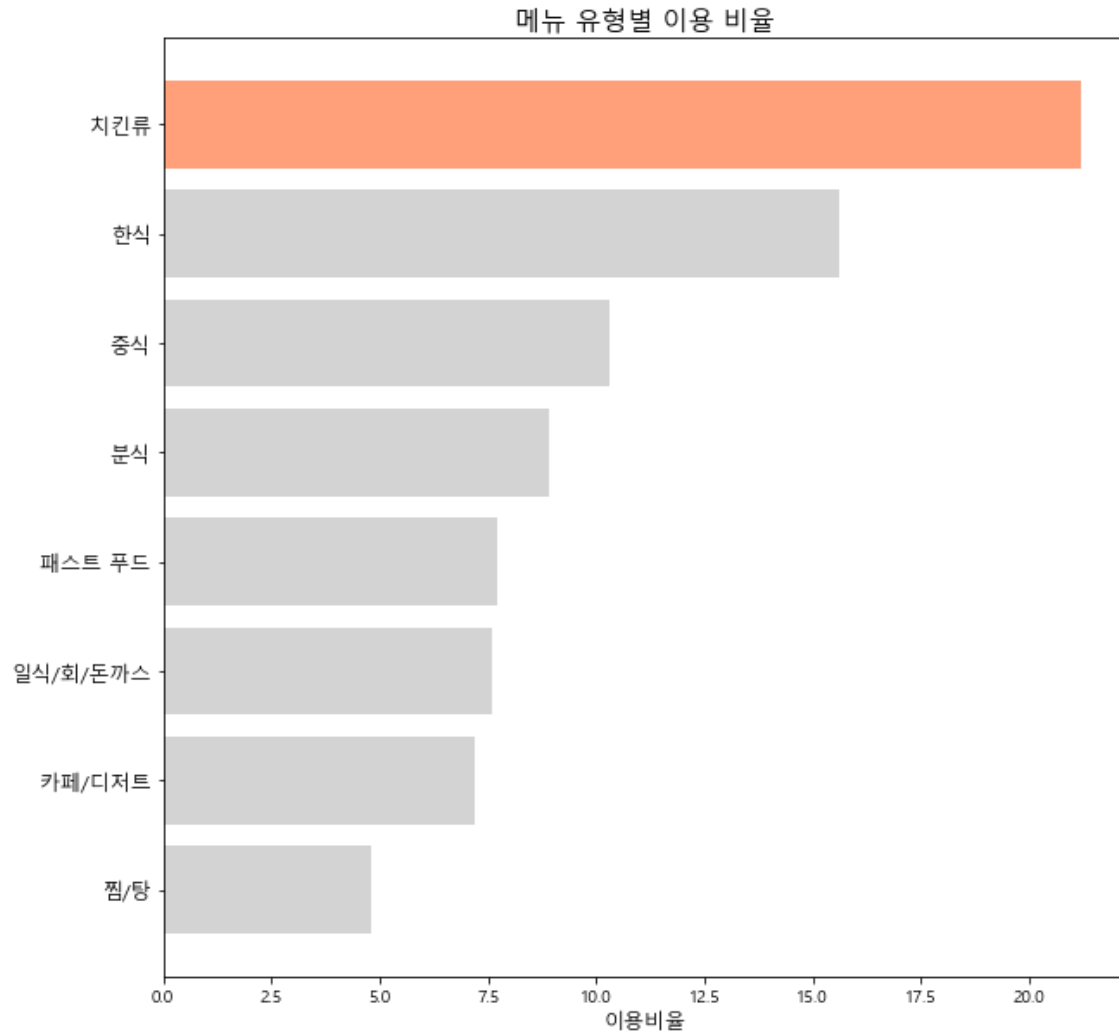
파스타계의 변화로는 **"생면"**이 떠오르고 있음

피자

식당을 언급하는 해시태그도 많았지만 **배달음식**으로 더욱 소비될 것



음식별(치킨) 분석



주요 배달 앱에서의 19년~ 21년
메뉴 이용 비율 순위

치킨 21.2 %

한식 15.6 %

중식 10.3 %

분식 8.9 %

패스트푸드 7.7%

“다섯 번 주문 중 한번은 치킨 주문”



음식별(치킨) 분석

태그 현황 분석

어떤 태그가 가장 많이 사용되었을까?

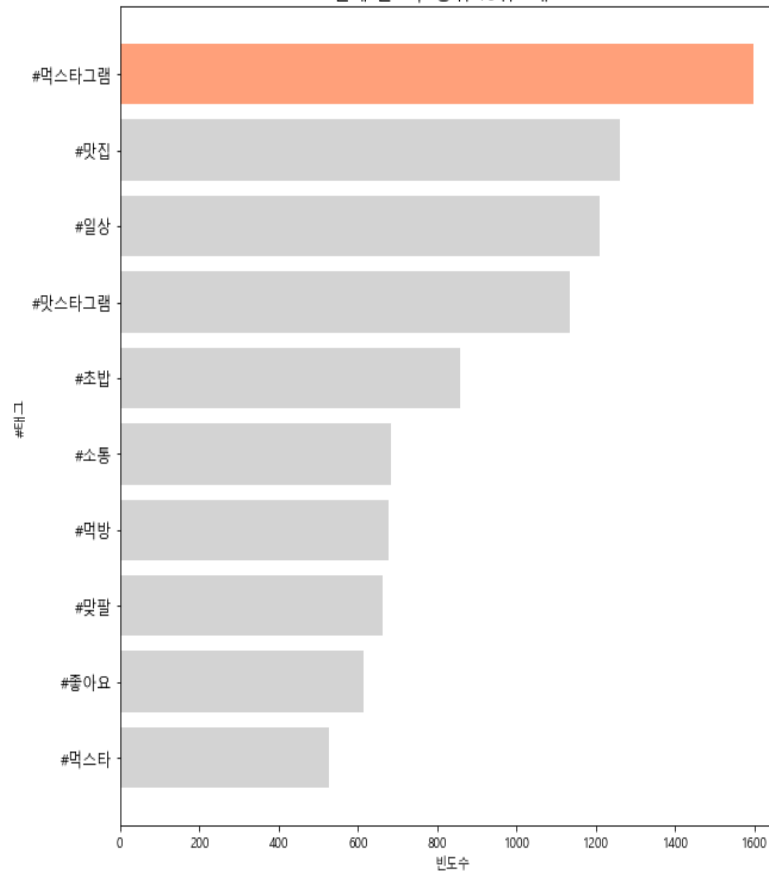
치킨 상위 10위

#먹스타그램 #맛집 #맛스타
그램 #일상

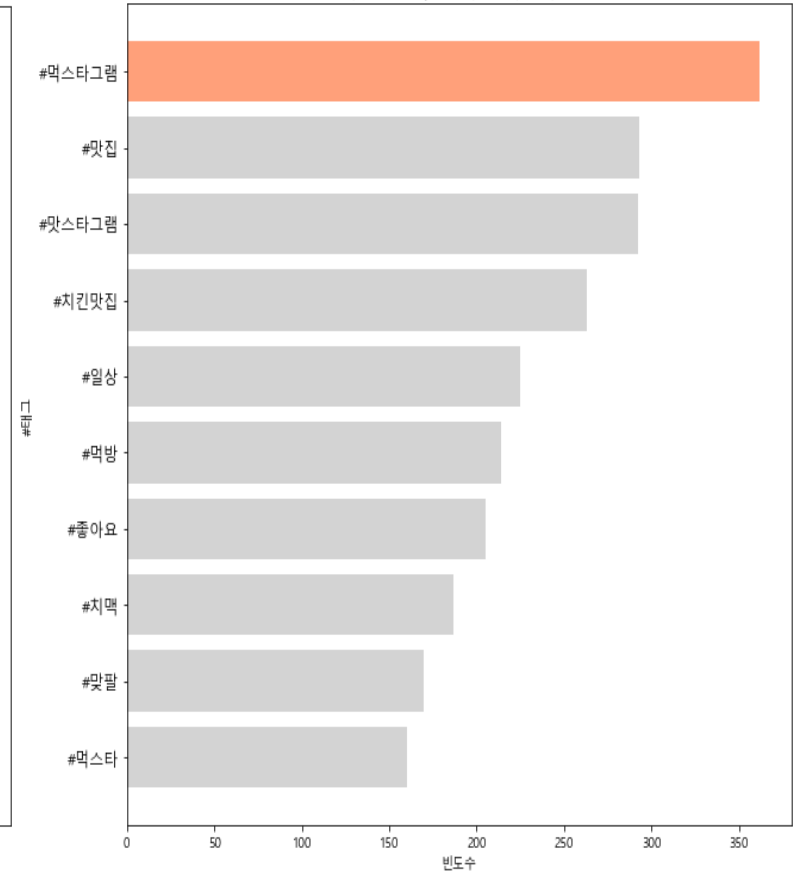
전체 음식 상위 10위

#먹스타그램 #맛집 #일상
#맛스타그램

전체 빈도수 상위 10위 #태그



치킨 빈도수 상위 10위 #태그





일상



일상 #일상스타그램 #먹스타그램 #맛팔 #좋아요

검색한 음식과는 직접 관련이 없는 계정 소유자들 간의 소통을 위한 키워드로 구성

광고 # 협찬



#야식치킨 #치맥 #신년회모임 #누구나홀딱반한닭

검색어와 직관적으로 연관이 있고 직접적인 브랜드명이나 페어링 메뉴 관련 키워드로 구성



일상

평균 좋아요 수 : 42.5

평균 본문 글자 수 : 182.2

평균 사용된 태그 수 : 22.2

평균 팔로워 수 : 1362.0



광고 # 협찬

평균 좋아요 수 : 74.2

평균 본문 글자 수 : 355.4

평균 사용된 태그 수 : 24.4

평균 팔로워 수 : 11146.4

#광고 #협찬 태그를 포함한 게시물의 **평균 좋아요 수**가 더 높다

광고 # 협찬 태그를 포함한 게시물의 **본문 글자 수**가 더 많다

#광고 #협찬 태그를 포함한 게시물의 **평균 사용된 태그 수**가 더 많다

광고 # 협찬 태그를 포함한 게시물의 **평균 팔로워 수**가 더 높다



“좋아요 수”가 높다?
불특정 다수에게 노출되기 쉬운 게시물!

“**좋아요**” 많이 받으려면?

팔로워 수

게시물 길이

게시물 수

+

“**엑스타그램**” 키워드 포함



팔로워 수

좋아요 상위 10프로 # 36103명
전체 평균 # 4888명

게시물 길이

좋아요 상위 10프로 # 292글
전체 평균 # 195글

게시물 수

좋아요 상위 10프로 # 1599개
전체 평균 # 969개



최종 추천 데이터 셋

Category	Plus_Tags	Essential_Tags
돈까스	['#돈까스', '#돈까스맛집', '#집밥', '#점심', '#치즈돈까스', '#food', '#우등', '#떡볶이', '#데일리', '#foodstagram']	['#먹스타그램', '#맛집', '#일상', '#맛스타그램', '#소통', '#먹방', '#맛팔', '#좋아요', '#먹스타', '#먹팔']
떡볶이	['#떡볶이', '#떡볶이맛집', '#분식', '#순대', '#집밥', '#koreanfood', '#튀김', '#김밥', '#야식', '#mukbangasmr']	['#먹스타그램', '#맛집', '#일상', '#맛스타그램', '#소통', '#먹방', '#맛팔', '#좋아요', '#먹스타', '#먹팔']
스시	['#스시', '#초밥', '#초밥맛집', '#사시미', '#서울맛집', '#일식', '#sushi', '#오마카세', '#회전초밥', '#초밥뷔페']	['#먹스타그램', '#맛집', '#일상', '#맛스타그램', '#소통', '#먹방', '#맛팔', '#좋아요', '#먹스타', '#먹팔']
치킨	['#치킨', '#치킨맛집', '#술스타그램', '#치킨추천', '#맥주', '#야식', '#치킨스타그램', '#food', '#데일리', '#일상스타그램']	['#먹스타그램', '#맛집', '#일상', '#맛스타그램', '#소통', '#먹방', '#맛팔', '#좋아요', '#먹스타', '#먹팔']
파스타	['#파스타', '#파스타맛집', '#스테이크', '#피자', '#샐러드', '#pasta', '#생면파스타', '#홍대맛집', '#연남동맛집', '#연남맛집']	['#먹스타그램', '#맛집', '#일상', '#맛스타그램', '#소통', '#먹방', '#맛팔', '#좋아요', '#먹스타', '#먹팔']
피자	['#피자', '#밥상일기', '#같이먹자', '#어린이집졸업', '#파스타', '#유치원입학', '#휴식시간', '#instafood', '#데일리', '#스파게티']	['#먹스타그램', '#맛집', '#일상', '#맛스타그램', '#소통', '#먹방', '#맛팔', '#좋아요', '#먹스타', '#먹팔']
햄버거	['#햄버거', '#수제버거', '#단체주문', '#잠실수제버거', '#단체도시락주문', '#감자튀김', '#food', '#맥도날드', '#점심', '#버거킹']	['#먹스타그램', '#맛집', '#일상', '#맛스타그램', '#소통', '#먹방', '#맛팔', '#좋아요', '#먹스타', '#먹팔']

Plus_Tags : 음식별로 나온 해시태그에서 Essential_Tags에서 검출되지 않은 **고유한 해시태그**

Essential_Tags : 전체 해시태그에서 **가장 빈번하게** 나온 해시태그



서론

X

본론

X

결론



https://finalproject/finale/결론

Instagram



홈



검색



탐색 탭



릴스



메시지



알림



만들기



프로필



더 보기

검색

검색

최근 검색 항목

모두 지우기

#

#1. 결론

X

1-1 구현

1-2 종합

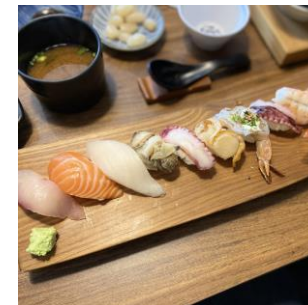
X

#

#2. 아쉬운 점

X

인기 게시물





구현

Instagram



홈



검색



탐색 탭



릴스



메시지



알림



만들기



프로필



더 보기

검색

검색

최근 검색 항목

모두 지우기

#

#1. 결론

X

인기 게시물

```
In [ ]: 1 img1 = 'data/chicken0.jpg'
        2 img2 = 'data/pork0.jpg'
        3 img3 = 'data/pasta0.jpg'
```

```
In [ ]: 1 img_list = [img1, img2, img3]
```

```
In [ ]: 1 plt.figure(figsize=(10,10))
        2
        3 for i in range(len(img_list)):
        4     img = img_list[i]
        5     img = cv2.imread(img, cv2.IMREAD_COLOR)
        6     img = cv2.resize(img, dsize=(128, 128))
        7     img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)
        8     img = img / 255.0
        9
        10 plt.subplot(1, 3, i+1)
        11 plt.title(img_list[i])
        12 plt.imshow(img)
        13 plt.axis('off')
        14 plt.tight_layout()
```

```
In [ ]: 1
```

```
In [ ]: 1
```

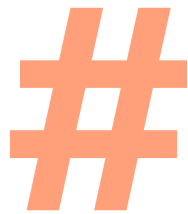
```
In [ ]: 1
```

```
In [ ]: 1
```



종합

결론



해시태그를 통해

언제 주로 먹을까?
어떤 음식과 자주 먹을까?
누구와 함께 먹을까?

LIKE를 많이 받기위해
팔로워 수, 게시물 길이, 게시물 수를 늘리고,
맛집 #먹스타그램 키워드를 추가하자!





아쉬운 점

아.쉽.다

해당 키워드와 관련없는 이미지 데이터도 함께 크롤링

인스타그램의 높은 보안수준으로 데이터 수집이 어려웠고, 많은 시간 소요

최종 프로그램 구현의 부재



How?

! 오답처리 ! 인스타그램_아이디 ! 실행가능_코드_구현



/Finale/Q&A



감사합니다!

Any Questions?