

Triple J가 분석한 2021년 여름 패션 트렌드

LikeLion 10회차 2팀
김지연 정희경 김현준



목차

1. 데이터 소개 및 분석 목적

1.1. 배경 및 주제 설명

1.2. 수집 데이터 설명

2. 데이터 수집 및 전처리

2.1. 수집 데이터 결과물

2.2. 전처리 과정 및 방법론

3. 데이터 분석 결과

3.1. 시각화 그래프와 인사이트

3.2. Challenge & Learning Point

1. 데이터 소개 및 분석 목적

- 주제 & 배경 설명
- 수집 데이터 설명

1. 데이터 소개 및 분석 목적

배경

- 셀레니움과 뷰티폴숲을 이용한 웹크롤링을 통해 데이터 수집부터 활용까지 실습
- 날씨와 패션을 공통 주제로 정하고 이와 관련된 데이터 수집

분석주제

- 미디어에서 언급되는 2021년 여름 패션 트렌드가 실제 소비자 구매로 이어지는지 파악



1. 데이터 소개 및 분석 목적

분석 목표

1. 미디어를 통한 2021년 여름 트렌드 조사

- 네이버 뉴스에 2021년 여름 트렌드 관련 키워드 검색을 통해 제품, 트렌드 파악
- 인스타그램에 #여름패션을 검색 트렌드 파악

2. 실제 소비자가 구매하는 제품 조사

- 국내 쇼핑몰 선정 후 베스트 아이템 데이터 수집

3. 1 과 2를 비교, 트렌드가 구매로 반영 되는지 검증



1. 데이터 소개 및 분석 목적

수집 데이터 설명

뉴스 기사

기간

2020. 2. 13 ~ 2021. 7. 5

용도

미디어에서 제안하는
2021년 여름 패션 트렌드 검색

방법

- 네이버 뉴스 웹 크롤링
- 데이터 전처리

인스타그램

기간

2021년 게시글 대상

용도

SNS에서 주로 언급되는 2021년
여름 패션 트렌드 탐색

방법

- #여름패션 해쉬태그게시물
- 웹 크롤링
- 데이터 전처리

온라인 쇼핑몰

기간

2021.6.26 ~ 2021.7.2

용도

온라인 쇼핑몰을 이용하는
소비자들의 구매 품목 파악

방법

- 온라인 몰 웹 크롤링
- 데이터 전처리

2. 데이터 수집 및 전처리

- 수집 데이터 설명
- 데이터 전처리 과정

2. 데이터 수집 및 전처리

데이터 수집 - How? (어떻게 수집하였는가?)

데이터 수집에 사용한 라이브러리



- 4개의 쇼핑몰 남여 베스트 아이템 400개
- Instagram 여름패션 게시물 1521개의 해쉬태그
- 네이버 여름패션 키워드 10페이지, 600건 기사

2. 데이터 수집 및 전처리

1) 쇼핑몰 데이터 수집 결과물

남성 & 여성 별 베스트 판매 패션 아이템 데이터

(각각 400개 총 800개)

추출 웹사이트 (4개)

서울스토어 / 코오롱몰 / SSF몰 / LF몰

	name	price	review	website
0	AP 스트레이트 데님 (Midnight Blue)	87200	420	Kolon
1	AP 1 턱 테이퍼드 슬렉스 (Just Black)	62300	71	Kolon
2	Relaxed Fit Logo Detail 포켓 티셔츠 (Natural)	34300	27	Kolon
3	AP Draw-String Easy 테이퍼드 슬렉스 (Just Black)	66750	38	Kolon
4	AP 1 턱 테이퍼드 슬렉스 (Mild Charcoal)	66750	71	Kolon
...

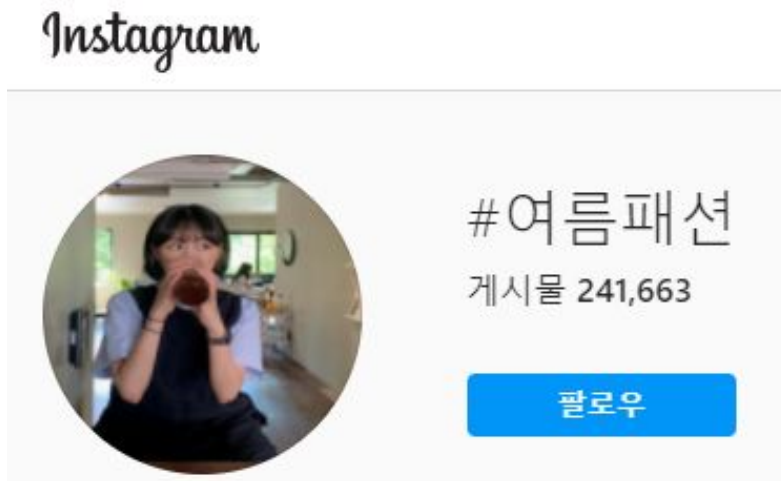


2. 데이터 수집 및 전처리

2) 인스타그램 데이터 수집 결과물

	좋아요	hashtag
0	1020	1or2 🍷
1	1020	플라워원피스
2	1020	베이직원피스
3	1134	협찬
4	1134	마크곤잘레스
...
8833	2009	천황산
8834	2009	등산
8835	2009	mountain
8836	2009	hiking
8837	2009	workout

여름 패션을 키워드로 검색한 게시물의 해시태그과 좋아요 숫자 데이터
(1521개 게시물에서 8837개 해시태그 추출)



2. 데이터 수집 및 전처리

3) 네이버 뉴스 데이터 수집 결과물

연문사	기사명	날짜	내용	키워드
0	조선일보 이준석과 봉태규 패션, 누가 몹쓸 핏?... MZ 세대 패션탐구 [23CM]	2021.06.14.	청취자들의 멋과 품격까지 신경쓰는 23CM가 이번 주 방송에서 2021년 여름 최...	2021년 여름 패션 트렌드
1	스냅 여름패션의 꽃 '프린트' 2021 베스트 트렌드	2021.06.22.	옷차림이 가벼워지는 여름패션은 하나의 아이템 만으로도 멋이 살아나는 프린트 스타일이 ...	2021년 여름 패션 트렌드
2	뉴스1 "패션·리빙·가전·식품 윙팡 최저가"...CJ온스타일 플러스, 6월 프로모션	2021.06.09.	2021년 여름 패션 트렌드인 플라워 프린팅에 여성스러운 실루엣을 더한 블라우스와 ...	2021년 여름 패션 트렌드
3	패션비즈 '집콕 패션' 패러다임 올 체인지	2021.06.09.	라이프'가 패션 트렌드까지 바뀌었다. 비대면 시대, 재택근무 등등 집에서 보내는 ...	2021년 여름 패션 트렌드
4	미디어 SR 2021년 '봄·여름' 패션 트렌드?... '클래식·시즌리스'	2021.01.15.	사진=이앤씨 팬데믹이 전환점이 될 2021년 봄·여름 패션 트렌드는 직장과 집을 ...	2021년 여름 패션 트렌드
...
595	시장경제신문 [시경Today-윙팡] 쿠팡, 가을 패션 브랜드 위크 외	2020.10.08.	패션위크는 쿠팡이 매달 첫 주 테마를 선정해 패션 트렌드 및 할인 상품을 제안하는 ...	2021년 패션 트렌드
596	글로벌예픽 어반테일러 수원점, 예비부부 위한 맞춤청장 프리미엄 서비스 제공	2021.03.11.	빛나는 결혼식을 완성시켜줄 결혼 예복은 물론 멋스러운 혼주 정장 등 맞춤형 예복을 ...	2021년 패션 트렌드
597	뉴데일리 [Daily New패션·뷰티] 루나, 아모레퍼시픽그룹, 신원 외	2020.09.21.	강화해 2021년 300억원, 2022년 400억원의 온라인 매출 달성을 목표로 하...	2021년 패션 트렌드
598	스냅 편안하고 스타일리시하다! 2021 봄 윙행 팬즈	2021.02.19.	2021년 봄·여름 패션 트렌드는 자연스럽게 표현되는 편안한 룩이 강세다. 정형화...	2021년 패션 트렌드
599	뉴스워커 신영와코루, BYC, 비비안 등 코로나19로 의류 시장 침체에도 '내의 업계'는 성...	2021-06-28	한편 '한국패션마켓트렌드 2021'에 따르면 내의류 전체 시장 규모의 70% 이상은...	2021년 패션 트렌드

2021년 패션 트렌드 검색어 뉴스 데이터

(총 600개의 뉴스 기사 데이터 수집)



2. 데이터 수집 및 전처리 전처리 과정

적용 1단계

모든 데이터 전처리

- 1) 자료형 변환
- 2) 공백문자 제거 (“\n”)
- 3) 불필요한 문자열 제거

적용 2단계

Text 데이터 전처리

- 1) 불용어 처리
- 2) Ngram() 함수 활용

적용 3단계

Feature Engineering

- 1) 기존 피처 결합
- 2) 피처 생성

2. 데이터 수집 및 전처리

전처리 과정 - 불용어 사전 작성 및 적용

```
hash_tag_stop_words = ['해쉬태그없음', '맛팔', '일상', '데일리', '줄반', '선팔', '좋아요', '감성', '육아', '인천환영',  
'뷰티톡', '홈카페', '러브', '더럽', '초보육아', '좋아요반사', '1or2', '소통', '선팔하면맛팔', '육아소통', '핵잼',  
'맛팔환영', '존맛탱', '줄튀', '대유잼', '맘스타그램', '초보맘소통', '육아맘', '몸짱', '육아스타그램', '익산',  
'익산그레이스', '익산옷가게', '군산옷가게', '전주옷가게', '대전옷가게', '광주옷가게', '택배', '익산맘', '워킹맘',  
'아들맘', '익산맛집', '익산커피숍', '예쁜커피숍', '프리랜서모델', '소통해요', '거울샷', '존맛', '줄테반사', '초보맘',  
'이벤트', 'fff', '분위기', '여름네일', '오오티디', '여름', '패피', '오늘', '패션', '광고', '여름패션', '여름코디',  
'여름신상', '여름옷', 'ootd']
```

```
after_stop_word = []  
  
for hash_tag in hash_tag_column:  
    if hash_tag not in hash_tag_stop_words:  
        after_stop_word.append(hash_tag)  
  
after_stop_word = nltk.Text(after_stop_word, name="해시태그")
```

- 1) 텍스트 데이터의 빈도수를 시각화
- 2) 확인 된 불용어를 리스트 자료화
- 3) 반복문을 활용하여 불용어 처리

2. 데이터 수집 및 전처리

전처리 과정 - Ngram() 함수

```
from nltk.util import ngrams
```

```
male_df["two_grams"] = male_df["item_list"].map(lambda x: list(ngrams((x), 2)))
```

	name	price	review	website	item_list	two_grams
0	AP 스트레이트 데님 (Midnight Blue)	87200	421	Kolon	[AP, 스트레이트, 데 님, Midnight, Blue]	[AP 스트레이트, 스트레이트 데님, 데님 Midnight, Midnight Blue]
1	AP 1 톱 테이퍼드 슬 렉스 (Just Black)	62300	72	Kolon	[AP, 1, 톱, 테이퍼드, 슬렉스, Just, Black]	[AP 1, 1 톱, 톱 테이퍼드, 테 이퍼드 슬렉스, 슬렉스 Just, Just B...
2	Relaxed Fit Logo Detail 포켓 티셔츠 (Natural)	34300	28	Kolon	[Relaxed, Fit, Logo, Detail, 포켓, 티셔츠, Natural]	[Relaxed Fit, Fit Logo, Logo Detail, Detail 포켓...
3	AP Draw-String Easy 테이퍼드 슬렉 스 (Just Black)	66750	39	Kolon	[AP, Draw-String, Easy, 테이퍼드, 슬렉 스, Just, Black]	[AP Draw-String, Draw- String Easy, Easy 테이퍼드, ...]

Ngram() 함수

결과값을 튜플
형태로 반환

단어 뭉치의
의미를 파악하기
위해 사용

2. 데이터 수집 및 전처리

전처리 과정 - 피쳐 결합

	name	price	review	website	item_list	two_grams	review_scaled	sales_index
0	AP 스트레이트 데님 (Midnight Blue)	87200	421	Kolon	[AP, 스트레이트, 데 님, Midnight, Blue]	[AP 스트레이트, 스트레이트 데님, 데님 Midnight, Midnight Blue]	0.004262	371.679947
1	AP 1 톱 테이퍼드 슬 렉스 (Just Black)	62300	72	Kolon	[AP, 1, 톱, 테이퍼드, 테 슬렉스, Just, Black]	[AP 1, 1 톱, 톱 테이퍼드, 테 이퍼드 슬렉스, 슬렉스 Just, Just B...	0.000729	45.414140
2	Relaxed Fit Logo Detail 포켓 티셔츠 (Natural)	34300	28	Kolon	[Relaxed, Fit, Logo, Detail, 포켓, 티셔츠, Natural]	[Relaxed Fit, Fit Logo, Logo Detail, Detail 포켓...	0.000283	9.723502

가격 * 리뷰 숫자 = 예상 Sales Index -> 기존 피쳐를 결합하여 생성

WHY?

아이템의 가격 차이 -> review 숫자에 영향을 줄 수 있기 때문
가격을 Sales Index의 가중치로 사용하기로 결정

2. 데이터 수집 및 전처리

전처리 과정 - 피처 생성

```
def one_mile_look(x):  
    look = ["슈트 팬츠", "긴팔 풀 오버", "스니커즈", "슈트 레깅스", "오버사이즈 스웨트 셔츠",  
            "슬리브리스 탑", "조거 팬츠", "스웨트 팬츠", "스웨트 셔츠", "바이커 슈트", "데님 슈트",  
            "크롭 티셔츠", "사이클링 슈트", "집업 후디", "폴로 셔츠", "오버사이즈 티셔츠"]  
  
    score = 0  
    for gram in x:  
        for one in look:  
            if gram == one:  
                score += 1
```

```
female_df["one_mile_look"] = female_df["two_grams"].apply(lambda x: one_mile_look(x))  
female_df["mun_sen_look"] = female_df["two_grams"].apply(lambda x: mun_sen_look(x))  
female_df["pokari_codi"] = female_df["two_grams"].apply(lambda x: pokari_codi(x))  
female_df["climbing_look"] = female_df["two_grams"].apply(lambda x: climbing_look(x))  
female_df["golf_look"] = female_df["two_grams"].apply(lambda x: golf_look(x))
```

패션 트렌드에 속하는
아이템 카테고리를
정성적 조사로 리스트
자료형으로 정리 및
점수 계산 함수 생성

웹사이트 데이터
Two_gram 피처
데이터에 함수 적용 및
아이템별 트렌드 피처
생성

2. 데이터 수집 및 전처리

전처리 과정 - 피처 생성

	name	price	review	website	two_grams	review_scaled	sales_index	one_mile_look	mun_sen_look
8	[24/7 series] 니트 탑 카라 셔츠 (COOLMAX) (247)	69000	515	Kolon	[24/7 series, series 니트, 니트 탑, 탑 카 라, 카라 셔츠, 셔츠...	0.005214	359.771593	0	0
9	[P3] 액티브 클 링 조거 팬츠 (냉감트리코 트)	109000	26	Kolon	[P3 액티브, 액 티브 클링, 클 링 조거, 조거 팬츠, 팬츠 냉 감트리코트]	0.000263	28.692632	1	0
14	[24/7 series] 니트 탑 카라 셔츠 (COOLMAX) (247)	69000	515	Kolon	[24/7 series, series 니트, 니트 탑, 탑 카 라, 카라 셔츠, 셔츠...	0.005214	359.771593	0	0

3. 데이터 분석 결과

- 시각화 그래프와 인사이트
- Challenge & Learning Point

3. 데이터 분석 결과

시각화 그래프와 인사이트



INSIGHT

인스타그램 사용자인 20-30대의 패션 키워드 파악
 특정 룩을 지칭하는 단어들을 해쉬태그로 소통
 코로나로 인해 등산 및 문화센터 활동과 연관된 패션 주목

국내 소셜 미디어 연령별 월평균 이용자 수 단위: 명

	10대	20대	30대	40대	50대
1위	 221만	 493만	 440만	 502만	 544만명

https://biz.chosun.com/site/data/html_dir/2020/09/07/2020090700514.html

3. 데이터 분석 결과

시각화 그래프와 인사이트

뉴스 데이터 워드클라우드

- 여름 패션 트렌드에도 코로나 영향이 지속됨
 - 편안함, 집콕, 집, 실용, 마일 (원마일룩), 일상, 외출, 활동
- 여름 휴가와 관련된 단어들 검색됨
 - 리조트, 휴가, 호텔, 수영복, 레저 등
- 여름과 관련된 색/ 2021년 관련된 색 추출
 - 블랙, 화이트, 블루, 팬톤
- 출시된 제품 및 인기 패턴에 대해 알아 볼 수 있음
 - 플라워, 꽃, 프린팅, 로고, 원피스, 티셔츠 팬츠, 블라우스, 패턴, 니트 등



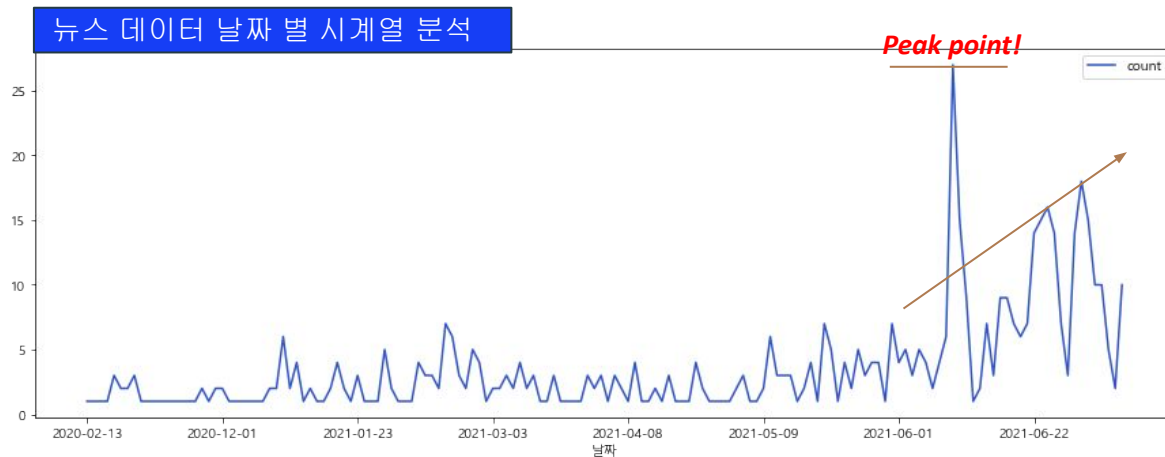
뉴스에서 언급된 트렌드 및 제품을 베스트아이템 리스트에 접목

3. 데이터 분석 결과

시각화 그래프와 인사이트

날짜	count
2021-06-09	27
2021-06-29	18
2021-06-24	16
2021-06-30	15
2021-06-23	15
...	...
2021-03-18	1
2021-03-02	1
2021-02-11	1
2021-02-10	1
2021-04-02	1

154 rows × 1 columns



INSIGHT

21년 6월부터 전체적으로 여름 패션 관련 뉴스 기사 발행이 증가

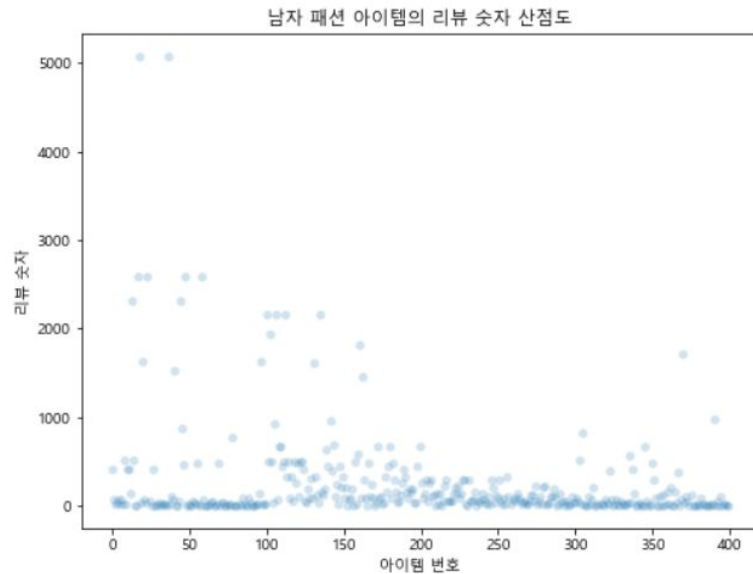
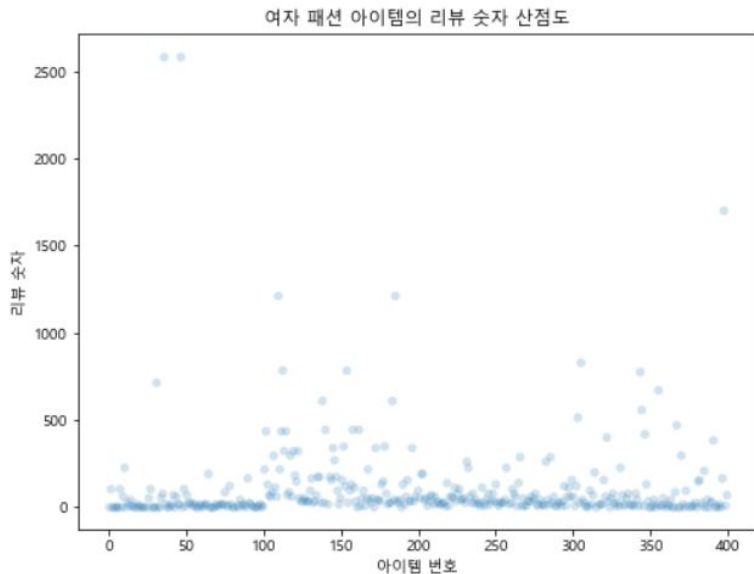
직관적으로 패션산업은 계절 민감성을 띄며, 계절 변화에 선행 하는 패턴이 있음
데이터로 직접 확인 할 수 있었음

3. 데이터 분석 결과

시각화 그래프와 인사이트

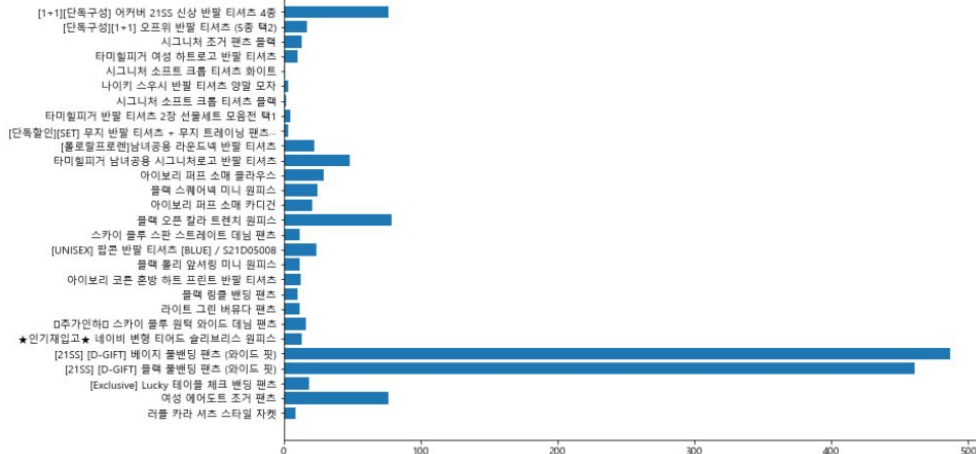
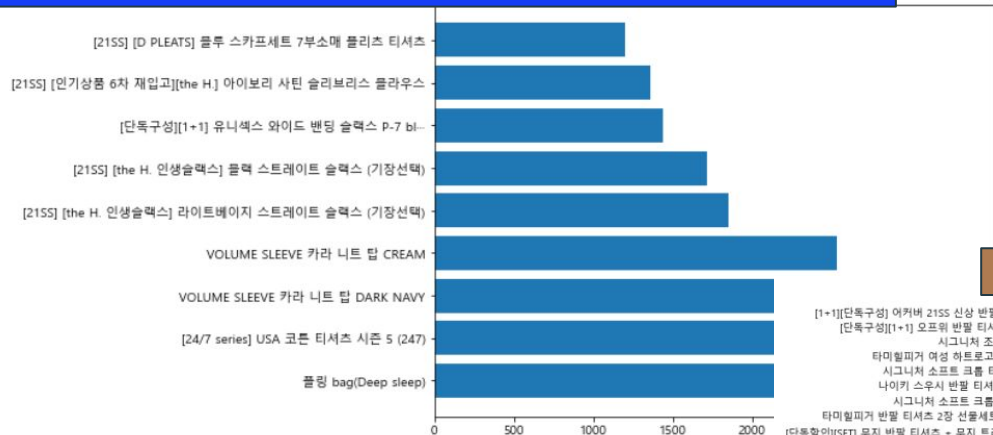
INSIGHT

베스트 아이템 목록에 속한 아이템 중에 **80%** 이상은 리뷰 숫자가 **500-1000**개 이하이다.
웹사이트의 베스트 아이템 선정 기준 <-> 방문자의 베스트 아이템에 대한 기대 (불일치 가능성)



3. 데이터 분석 결과 시각화 그래프와 인사이트

Sales Index 기준 여성 패션 상위 10개 아이템



INSIGHT

두개의 리스트에 겹치는 아이템? 없음

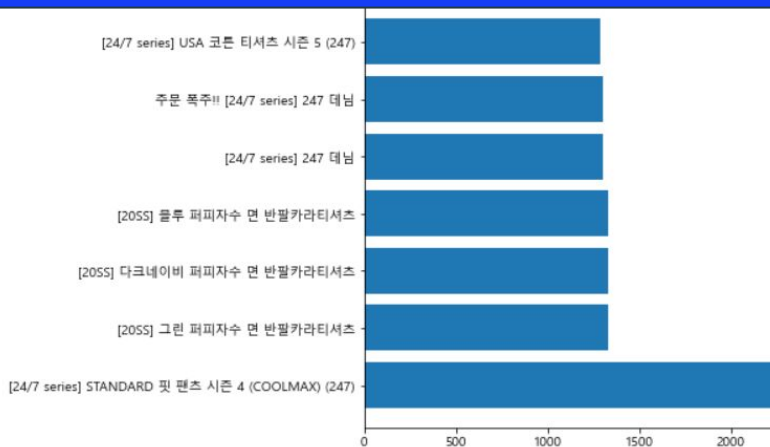
- 웹사이트의 베스트 세일즈 아이템

- 뉴스 & 인스타 패션 트렌드
아이템 간의 상관관계가 낮음

3. 데이터 분석 결과

시각화 그래프와 인사이트

Sales Index 기준 남성 패션 상위 10개 아이템



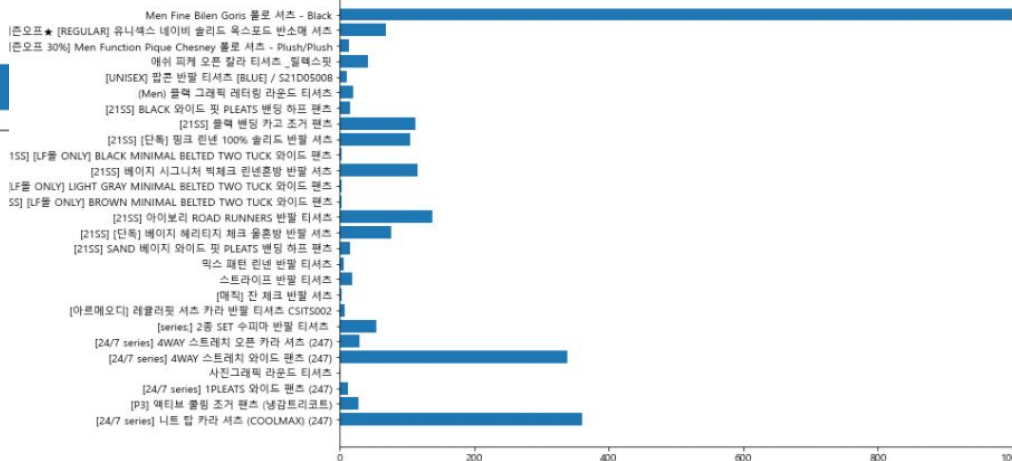
INSIGHT

두개의 리스트에 겹치는 아이템? 없음

- 웹사이트의 베스트 세일즈 아이템

- 뉴스 & 인스타 패션 트렌드
아이템 간의 상관관계가 낮음

Fashion Trend 기준 남성 패션 아이템 목록



3. 데이터 분석 결과

Challenge & Learning Point 1: 텍스트 분석의 한계

- 여러가지 단어로 한 가지 아이템을 명명하는 케이스 ex) 반팔 티셔츠



amoupre
originals

J3TAS21501KH-X

[올데이 티셔츠] 흡습속건

Relaxed Fit Logo Detail Pocket T-Shirt (Natural)



SEARCH410
selected

K16238057204030679HAKJ

[UNISEX] FORMULA LOGO OVER T



GGIO2
originals

AXWAM20422BUX

톤온톤 스트라이프 반소매 니트



CUSTOMELLOW
originals

CWTAM21483C

[백향목 X 커스텀멜로우] "LOVE ONLY" short sleeve t-shirt



SERIES
originals

SATAW21901OTX

[series:] 2종 SET 수피마 반팔 티셔츠

원피스 맛집으로 통하는 로맨틱 페민닌 브랜드 '샤틴(SATIN)'은 화려한 **플라워** 패턴 원피스, 복고풍 도트 무늬 원피스, 여성스러운 머메이드 원피스, 지적인 플리츠 원피스 등 봄과 여름까지 한번에 책임지는 매력적인 원피스로 여성 고객들의 마음을 사로잡



STYLE

SNEAKERS

GROOMING

ENTERTAINMENT

TRAVEL & EATS

CAR & TECH

WATCH

VIDEO

SEX

STYLE

2021.06.30

2021 S/S 룩을 입고 즐기는 영화적 여름

2021 S/S 룩을 입고 영화적 여름을 보낸다.

Romeo + Juliet 15세기 이탈리아 베로나에서 1990년대 미국 마이애미로 배경을 옮긴 1996년 버전 <로미오 + 줄리엣>. 영티에서 촬영해 이국적이고 뜨거운 태양의 열기로 가득 차 있다. 배경은 바뀌었지만 몬터규가와 캐폴렛 가문의 대로 가져와 배우들의 의상에도 이를 반영했다. 로미오와 그의 사촌들은 모두 화려한 패턴의 하와이안 셔츠를, 색의 가족 옷을 입고 등장한다. 그래서 레오나르도 디카프리오의 미모만큼 그가 입고 나오는 셔츠는 인상적이다. | 존 팜 엔젤스의 런웨이에서도 로미오가 입었던 셔츠와 꼭 닮은 일본풍 **플라워** 패턴 셔츠가 등장했다. 야자수 가카, 팝한 컬러의 베르사체도 마찬가지. 이국적인 하와이안 셔츠 하나면 영화 속 해변으로 데려가줄 듯하다.

[Consumer Journal] **꽃무늬** 할매룩, 꾸안꾸 남친룩

"내 취향이 트렌드죠" 유행보다 익스앤드매처로 개성 입는다

올해의 색 노랑 기본으로

밝은 회색이나 보색과

맞춰 입으면 좋아

남녀 레트로풍 인기 여전

리넨소재로 편안한 느낌 연출

이영욱 기자 | 입력 : 2021.05.06 04:02:05

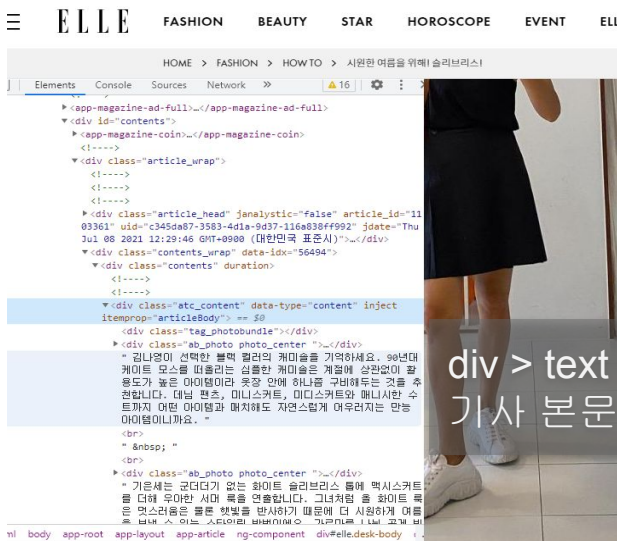
0

뉴스, 매거진 기사도 국문 및 영문 혼용해서 사용 → 패션 관련 분석은 이미지데이터 활용이 효과적

3. 데이터 분석 결과

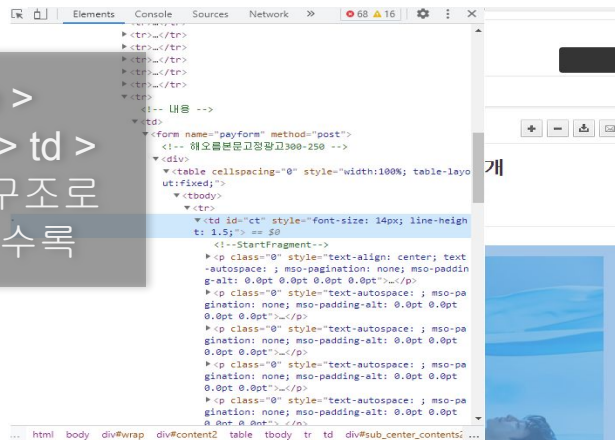
Challenge & Learning Point 2: 크롤링과 코딩의 어려움

뉴스를 제공하는 사이트별 html 구조가 상이하여 기사 전문을 가져오는 데 한계가 있었음



div > table >
tbody > tr > td >
p > span 구조로
기사 본문 수록

div > text 구조로
기사 본문 수록



[패션저널&텍스타일라이프:원유진 기자] 에프앤에프(F&F대표 김창수)의 프리미엄 라이프스타일 아웃도어 브랜드 '디스커버리 익스페디션'이 본격적인 휴가철을 맞아 활동적인 액티비티부터 여유로운 리조트에서의 휴까지 즐길 수 있는 다양한 스타일의 한서머 룩을 공개했다.

이번 시즌에는 현실과 상상의 경계를 넘나드는 디스커버리의 시원한 여정을 판타지를 주제로 미지의 세계를 탐하는 공룡의 모습을 감각적인 비주얼로 담아냈다.

3. 데이터 분석 결과

Challenge & Learning Point 2: 크롤링과 코딩의 어려움

기사 작성 날짜 양식에도 차이가 있어 통일시키는 작업을 진행해야 했음

→ 대규모 데이터를 가져오려면 **html** 코드와 **Selenium**과 **BeautifulSoup**에 대한 이해 & 고성능의 서버와 컴퓨팅 파워가 필요

```
In [72]: data['날짜'].unique()

Out[72]: array(['2021.06.14', '2021.06.22', '2021.06.08', '2021.01.15',
                '11면 1단', '2021.04.26', '2021.04.09', '2021.04.01',
                '2021.04.06', '2021.02.25', '2021.02.15', '8면 1단',
                '2021.06.10', '2021.02.27', '2021.02.23', '2021.03.04',
                '2021.03.12', 'A20면 1단', '2021.01.23', '2021.03.08',
                '2021.03.03', '2021.06.23', '2021.01.18', '2021.02.01',
                '2021.01.06', '2021.06.25', '2020.10.22', '2021.06.24',
                '2021.02.19', '2020.07.16', '2021.06.11', '2021.06.05',
                '2019.06.18', '2018.05.10', '2021.06.02', '2021.06.04',
                'C1면 TOP', '2021.06.01', '2021.05.14', '2021.04.20', '8면 1단',
                '2021.04.23', '2021.03.22', '2021.04.07', '2021.04.14',
                '06면 1단', '2021.02.14', '14면 1단', '2021.03.17', '2021.03.16',
                '2021.03.02', '2021.02.24', '2021.02.17', '2021.02.16',
                '19면 TOP', 'C7면 TOP', '2020.12.01', '2020.11.29', '2021.01.04',
                '2021.02.18', '2020.07.17', '2020.06.05', '2019.06.15',
                '2021.06.17', '2021-07-05', '2021-07-02', '2021.06.26',
                '2021-06-28', '2021-06-29', '2021-07-01', '2021-06-30',
                '2021.06.16', '2021.06.08', 'A12면 4단', '2021.06.07',
                '2020.12.11', 'B4면 1단', '21면 TOP', '2021.05.31', '2021.05.26',
                '2021.03.30', '2021.05.11', '2021.05.10', '2021.06.03',
                '2021.05.25', '2021.05.30', '2021.05.28', '2021.05.27',
                '2021.05.13', '2021.05.19', '2021.05.23', '2021.05.24',
                '2021.04.08', '2021.05.17', '2021.05.20', '2021.05.18',
                '2021.06.21', '13면 1단', '2021.05.04', '2021.04.02',
                '2021.04.27', '2021.04.25', '2021.03.11', '2021.06.27',
                '2021.03.31', '2021.03.23', 'C3면 TOP', '2021.01.26', '15면 TOP',
                '2020.12.31', '2020.09.08', '2020.11.05', '2021.01.14',
                '2020.09.11', '2020.04.03', 'B11면 1단', '2020.08.14',
                '2020.07.20', '2020.03.30', '2020.05.07', '2020.02.13',
                '2021.06.06', '2021.05.07', 'A31면 1단', '2021.04.28',
                '2021.05.01', '2021.05.21', '2021.05.03', '2021.04.22',
                '2021.04.16', '2021.05.16', '2021.05.09', '2021.04.12',
                '25면 TOP', '2021.05.06', '16면 3단', '2021.06.19', '2021.06.18',
                '2021.06.15', '2021.06.13', '2021.05.22', '2021.03.18',
                '2021.04.23', '2021.03.19', 'A17면 1단', '2021.05.02',
                '2021-07-04', '2020.09.23', '2021.01.29', '2021.06.12',
```

```
In [80]: data_re = []
for date in data['날짜']:
    new_date = date[0:10]
    new_date2 = new_date.replace('.', '-')
    print(new_date2)
    data_re.append(new_date2)
print(data_re)
print(len(data_re))
```

2021-06-14
2021-06-22
2021-06-09
2021-06-09
2021-01-15
2021-04-26
2021-04-09
2021-04-01
2021-04-06
2021-02-25
2021-02-15
2021-06-09
2021-06-10
2021-02-27
2021-02-23
2021-03-04
2021-06-09
2021-03-12
2021-01-23
2021-03-08

‘XXXX-XX-XX’, ‘XXXX.XX.XX.’ 또는
날짜 이외의 문자가 섞여 있어서
날짜 이외의 문자를 삭제하고
날짜는 ‘XXXX-XX-XX’의 한 가지
형식으로 통일

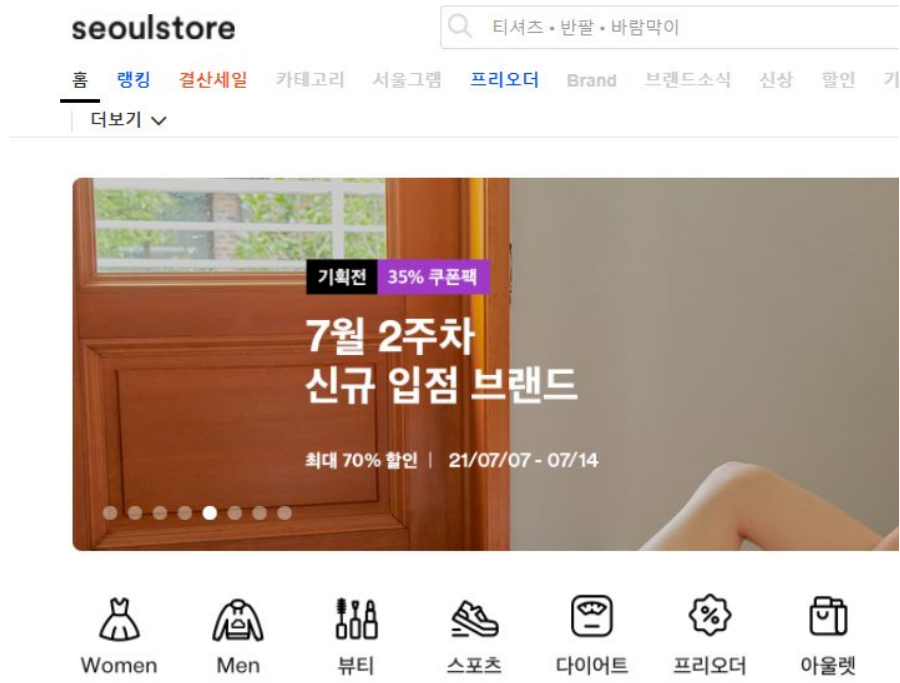
```
In [81]: data
```

3. 데이터 분석 결과

Challenge & Learning Point 3: 도메인 지식 & 데이터 수집의 한계

패션 산업의 트렌드 파악 → 배경지식이 필요

트렌드를 반영할 판매 데이터 → 내부 데이터



감사합니다



출처

- <https://www.marieclairekorea.com/fashion/2021/01/2021-ss-trend-report/>
- https://www.fashionn.com/board/read_new.php?table=1004&number=36662
- https://www.gqkorea.co.kr/2021/06/30/2021-s-s-%eb%a3%a9%ec%9d%84-%ec%9e%85%ea%b3%a0-%ec%a6%90%ea%b8%b0%eb%8a%94-%ec%98%81%ed%99%94%ec%a0%81-%ec%97%ac%eb%a6%84/?utm_source=naver&utm_medium=partnership
- <https://www.mk.co.kr/news/economy/view/2021/05/434487/>
- <https://www.kolonmall.com/>
- https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.youtube.com%2Fwatch%3Fv%3DxLuizC2AQgk&psig=AOvVaw1-123zsn7qlzoOzjZkgTfK&ust=1625806702015000&source=images&cd=vfe&ved=0CAkQ3YkBa_hcKEwilhtCM2NLxAhUAAAAAHQAAAAAQAw
- https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.everypixel.com%2Fimage-9470256167309010299&psig=AOvVaw16WDaD4luLt7WBXPdpiAV_&ust=1625812241137000&source=images&cd=vfe&ved=0CAkQ3YkBahcKEwiY4Jne7NLxAhUAAAAAHQAAAAAQMQ
- <https://www.elle.co.kr/article/56494>
- <http://okfashion.co.kr/detail.php?number=56558&thread=81r06r01>