

『네이버데이터랩』및『SNS데이터』를활용한 "FRESHEASY 밀키트" 분석

- Based on HMR market analysis

구자미 김수연 손예선 한지예

목차

분석 배경 및 목적

2. 데이터 수집 방법

3. 데이터 분석 방법

4. 데이터 전처리 과정

5. 분석결과 및 시각화

현재 환경과 HMR시장 중심

여러 경로에서 여러 데이터 수집

인지도, 워드 클라우드, 태블로, 상관관계, 시계열

Konlpy, WordCloud, fbProphet 등 사용

분석결과를 시각화

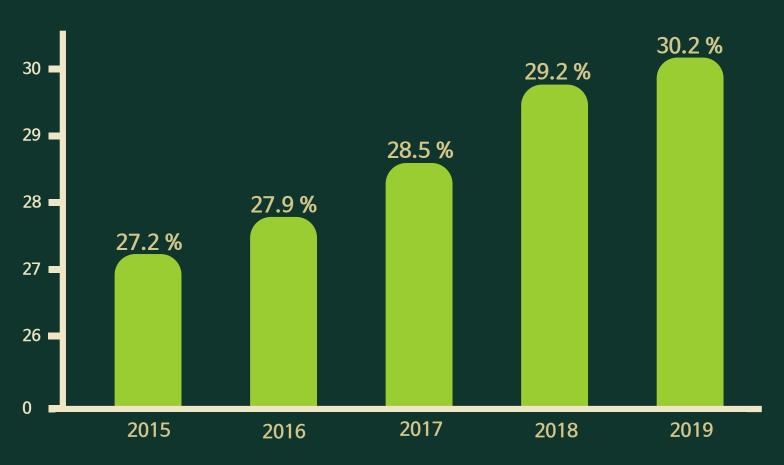


1. 분석 배경 및 목적

1. 분석 배경 및 목적

점점 증가하는 1인 가구 비율

자료출처: 통계청



가성비 가심비 가시비

HMR 이란?

Home Meal Replacement 의 약자 일반적인 식사 준비과정에서의 시간과 노력을 줄이려는 목적에서 탄생 어느정도 조리가 된 상태에서 가공,포장되어 단순한 조리과정으로도 음식이 완성되는 형태



RTP (식재료)

Ready to Prepare

세척후 소포장된 식료품

Ex) 씻어나온 쌀, 다진마늘, 소분된 파



RTC (반조리)

Ready to Cook

바로 조리가 가능하도록 손질된 제품

Ex) 밀키트



RTH (완조리)

Ready to Heat

데우면 바로 먹을 수 있는 제품 Fv) 주서떠보이

Ex) 즉석떡볶이, 레토르트 식품



RTE (즉석조리)

Ready to Eat

포장만 제거하고 바로 먹을 수 있는 제품

Ex) 도시락, 삼각김밥

밀키트란?

Meal(음식) + Kit(필요한 것을 모아 놓은 세트) 손질된 식재료와 딱 맞는 양의 양념, 조리법을 세트로 구성해 제공하는 제품 조리 직전의 냉장상태로 배송되며, 소비자가 직접 요리하는 형식

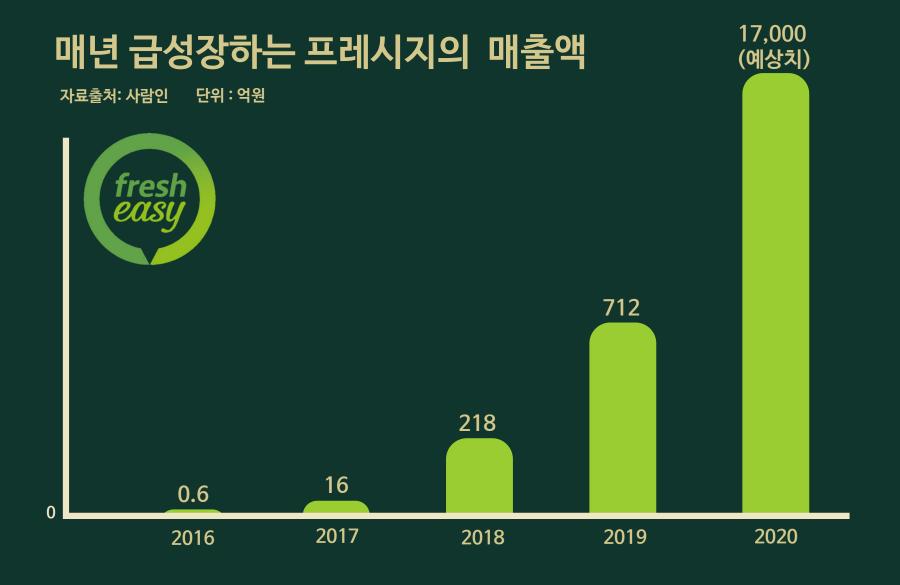




HMR시장과 밀키트 시장의 규모

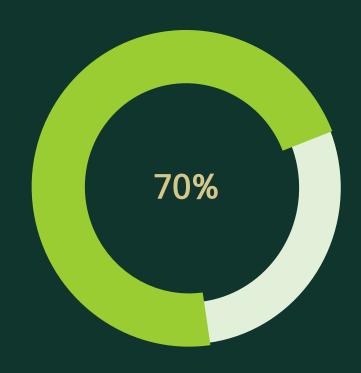
자료출처: 한국농촌경제연구원





국내 밀키트 시장 프레시지의 시장점유율





현재 판매중인 자체브랜드 제품 약 134 종

홈쇼핑, 온라인판매, 백화점에도 입점해 판매경로 확장 중



우삼겹치즈 쫄면 12,900원 12/31 수령 주문가능



알꼬막 깍두기 볶음밥(간장맛) 3,500원

12/31 수령 주문가능



[당일배송] the EASY 청주식 돈짜글 6,900원

당일배송 12/26(내일) 오후 11시 30문까지 도착



파빌리온 설악 황태 진국 14,900원

12/31 수령 주문가능



블로네제 라구 파스타 13,900원

12/31 수령 주문가능



갈릭가득 바게트볼 20%

2,000원 2,500원 12/31 수령 주문가능



the EASY 6종 SET(광양식소불고기+동경식규동+춘천 식양념닭칼비+서울식간장닭칼비+단짠제육볶음+매콤 제육볶음)

7% 38,900원 41,400원



[당일배송] the EASY 서울식 간장 닭갈비 6,900원

당일배송 12/26(내일) 오후 11시 30분까지 도착

국내 밀키트 기업 중 유일한 대량생산 시스템

올해 용인에 700억원을 투자하여 약 8천평 규모의 HMR 공장 설립



등장이후 급성장한 원인에 대한 예측

- 1. 가구의 형태 변화
- 2. 생활양식의 변화
- 3. 올해 초 등장한 COVID-19





2. 데이터 수집 방법

크롤링

- **뉴스 제목** (코로나 외식 관련 뉴스)
- **인스타그램** (밀키트 등 관련 키워드)
- **네이버 블로그** (밀키트 등 관련 키워드)
- **네이버 카페** (밀키트, 쿠킹박스 등)

기존 데이터 활용

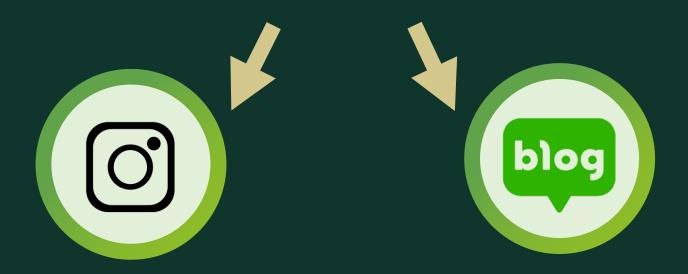
- 통계청 자료 다운로드
 (1인 가구 비율)
- **네이버 데이터랩 활용** (검색량 데이터)
- **농촌경제진흥연구원** (밀키트 판매량)
- 금융감독원 전자공시 시스템 (매출추이)

1) 크롤링

목적 1: 프레시지 인지도 분석

→ SNS 내 관련 키워드 중 프레시지가 차지하는 비율

밀키트, 밀키트추천, 홈파티음식, 홈파티레시피



1) 크롤링

목적 2: 프레시지 관련 키워드 분석

프레시지, 프레시지밀키트, 프레시지쿠킹박스



1) 크롤링

목적 3: 프레시지 리뷰 키워드 분석

네이버 프레시지 밀키트 상품 리뷰



2. 데이터 수집 방법

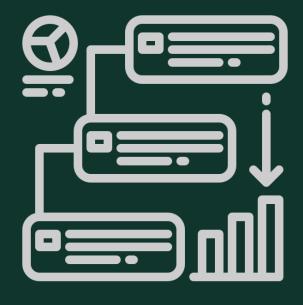
2) 기존 데이터 활용

목적: 키워드들의 상관분석과 향후 3개월간의 검색량 예측

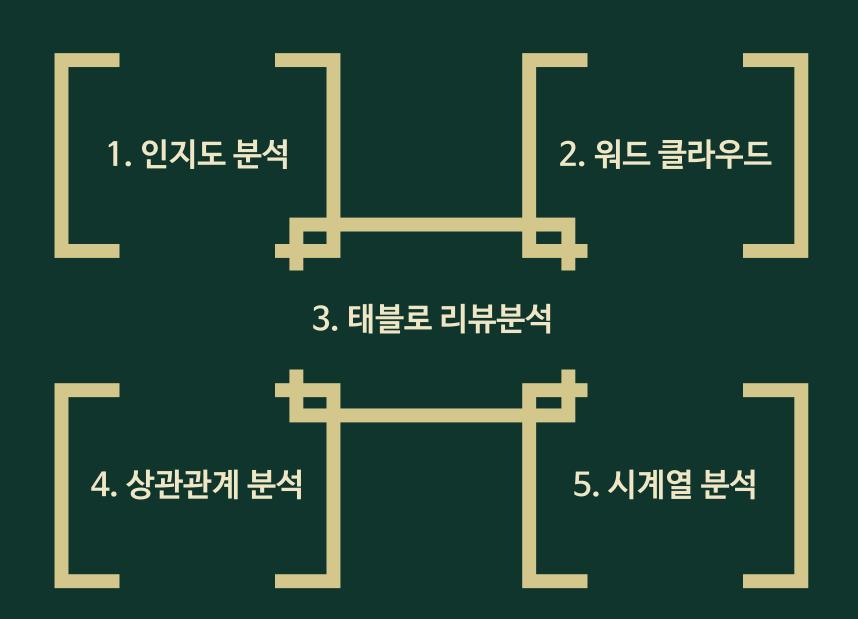
• **통계청** → 가구 형태별 총계

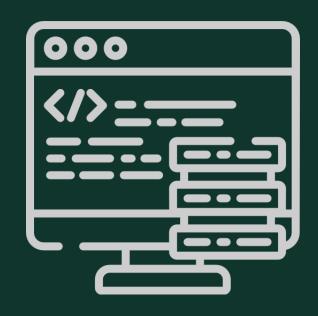
• 네이버 데이터랩 → 키워드

키워드: 프레시지, 간편식, 쿠킹박스, 밀키트, 레시피박스, 쉬운요리



3. 데이터 분석 방법





4. 데이터 전처리 과정

4. 데이터 전처리 과정

1) 인지도 분석

수집한 자료의 총 개수 확인

blog_l	kn		
	title	content	date
0	밀키트, 피코크 마라샹궈 괜찮아요	요즘 제가 꽂힌 #밀키트 이마트 피코크 마라샹궈에요ㅎ 마라탕은 못 먹는데 마라샹	2020.11.26.
1	국물떡볶이 은영이 떡볶이 캠핑 밀키트도 군!	캠핑 밀키트도 굳! 아이들 간식으로도 인기만점이지만, 사실 어른들 이 더 좋아하	2020.11.03.
2	경산 간편조리 밀키트 등갈비찜 & 해물누룽지탕 굿	밀키트 정말 잘 나오는군요 ㅎㅎ 그저 레시피대로 따라하면 됩니다 ㅎㅎ 만들고	2020.11.22.
3	부대찌개 밀키트 간편하고 맛있는 창영이 부대찌개	때 밀키트 요리를 많이 이용합니다. 모든 재료가 손질되어 있어 넣고 끓이면 완	5일 전
4	홈파티음식 잇츠온 밀키트로 양식 한상 뚝딱 준비 해요	홈파티음식 잇츠온 밀키트로 양식 한상 뚝딱 준비해요 벌써 올 한해 가 홀~쩍	3일 전
	iii.		2000
3995	잇츠온 양식 밀키트 홈파티 세트로 미리 크리스마 스 파티!	맛있게 파티음식을 차릴 수 있다면 자주 이용할 의사가 있어요. 파티 음식을	3일 전
3996	※집에서도 예쁘게 홈파티음식 만들기!	해서 홈 파티를 하게 되었습니다. 홈 파티 음식 추천! 홈 파티 음식물 준비하	2020.09.03.
3997	(6) : 어스, 핸드위치 US, HANDWICH 배달 / 나홀로 홈파티 무드	연말인데 금요일인데 혼자 있으려니 적적해서 혼자 홈파티를 해보기 로 한다. 파티	어제
3998	광주 화정동 맛집 홈파티 음식으로 테스팅노트	미리 전화로 12시까지 찾으러 간다 예약하고 갔었는데 딱 맞게 나와 서 음식이	2020.09.18.
3999	홈파티음식 핑거푸드 카나페만들기	안녕하세요~ 한글날도 지나고 벌써 월요일이네요^^ 오늘은 홈파티 음식 크랩카나페	2020.10.12.
4000 r	ws × 3 columns		

1) 인지도 분석

정제한 데이터 중 프레시지를 포함한 행의 개수 확인

```
blog_kn.drop_duplicates(subset = ["title"], inplace = True)
print(blog_kn)
blog_kn.to_excel('.../data/crawling_blog_kn.xlsx', index = False)
                                    title #
                         밀키트, 피코크 마라샹궈 괜찮아요
                    국물떡볶이 은영이 떡볶이 캠핑 밀키트도 굳!
                 경산 간편조리 밀키트 등갈비찜 & 해물누룽지랑 굿
부대찌개 밀키트 간편하고 맛있는 창영이 부대찌개
홈파티음식 잇츠온 밀키트로 양식 한상 뚝딱 준비해요
              연말 홈파티 음식에 진심인 사용원에서 간편하게 준비 끝!!
                        ※집에서도 예쁘게 홈파티음식 만들기!
3996
    (6) : 머스, 핸드위치 US, HANDWICH 배달 / 나홀로 홈파티 무드
광주 화정동 맛집 홈파티 음식으로 테스팅노트
3998
                          홈파티움식 핑거푸드 카나페만들기
3999
                                       content
                                                     date
       요즘 제가 꽂힌 #밀키트 이마트 피코크 마라샹궈에요ㅎ 마라탕은 못먹는데 마라샹...
1.26.
       캠핑 밀키트도 굳! 아이들 간식으로도 인기만점이지만, 사실 어른들이 더 좋아하... 2020.1
1.03.
       밀키트 정말 잘 나오는군요 ㅎㅎ 그저 레시피대로 따라하면 됩니다 ㅎㅎ 만들고 ... 2020.1
1.22
len(blog_kn)
3593
```

```
fresheasy = blog_kn['content'].str.contains('프레시지|fresheasy|Fresheasy|FRESHEASY|프레시이지')
fresh_count = blog_kn[fresheasy]
len(fresh_count)
```

2) 워드 클라우드

같은 방법으로 수집한 데이터 중 중복자료, 불순물 제거

```
STOPWORDS = ['#프레시지', '#fresheasy', '#프레시지밀키트', '#밀키트', '', '#프레시지쿠킹박스',
           '#맞팔', '#협찬', '#쿠킹박스', '#소통', '#좋아요반사', '#일상스타그램', '#선팔하면맞팔',
          '#선팔', '#좋반', '#일상그램', '#daily', '#데일리', '#음식', '#hi.hi.hi_mom', '#대현',
          '#먹스타맞팔', '#먹팔', '#선팔맞팔', '#일상', '#좋아요', '#럽스타그램', '#정대현', '#BAP',
          '#2in1', '#먹스타그램', '#맛스타그램']
tag_total_selected = []
for tag in tags_total:
   if tag not in STOPWORDS:
      tag_total_selected.append(tag)
tag counts selected = Counter(tag total selected)
tag counts selected.most common(50)
[('#집밥', 300),
('#스테이크', 270),
 ('#홈쿡', 206),
 ('#요리스타그램', 200),
('#밀푀유나베', 183),
 ('#블랙라벨스테이크', 142),
 ('#감바스', 136).
('#온더테이블', 122),
('#요리', 115),
('#미상터시티', 104),
```

2) 워드 클라우드

정제한 데이터를 토대로 워드 클라우드 생성

3) 태블로 리뷰 분석

네이버 리뷰에서 수집한 데이터 개수 확인 후 중복 데이터 삭제

```
review = pd.concat([steak_df, nabe_df])
len(review)

19943

review.drop_duplicates(subset='content', inplace=True)
len(review)

19867
```

3) 태블로 리뷰 분석

바른 말로 수정 & 의미 없는 데이터 삭제(1글자)

```
num = [index for index, value in enumerate(count_pos_selected) if value == '빨르다']
for i in range(len(num)):
   count_pos_selected[num[i]] = count_pos_selected[num[i]].replace(count_pos_selected[num[i]], '배르다')
pos_list = []
for i in range(len(review)):
    pos = t.pos(review['content'][i], norm=True, stem=True)
    pos list.extend(pos)
                                             for i in range(len(pos_list)):
                                                  if len(pos_list[i][0]) == 1:
len(pos_list)
                                                       del pos_list[i]
335151
                                              len(pos_list)
                                             221890
```

3) 태블로 리뷰 분석

동사와 형용사만 추출

```
pos_selected = []
for i in range(len(pos_list)):
    if pos_list[i][1] in ['Verb', 'Adjective']:
        pos_selected.append(pos_list[i][0])

len(pos_selected)

102506
```

```
count_pos = Counter(count_pos_selected)
count_pos.most_common(50)
[('맛있다', 9160),
 ('좋다', 7554),
 ('간편하다', 1840),
 ('편하다', 1625).
 ('신선하다', 1142),
 ('괜찮다', 1063),
 ('편리하다', 886),
('빠르다', 733),
 ('만족하다', 729),
 ('많다', 725),
 ('좋아하다', 665),
 ('부드럽다', 665),
 ('저렴하다', 650),
 ('깔끔하다', 616),
('꼼꼼하다', 533),
 ('적다', 531),
 ('간단하다', 513),
('아쉽다', 498),
 ('성성하다', 475),
 ('감사하다', 364),
 ('부족하다', 309),
```

4) 상관관계 분석

데이터 랩에서 수집한 일별 키워드 검색량 데이터로 히트맵 그래프 생성

result12.corr()							
		7L记시		쿠킹박스	ALOOJ	게 나파바스	ᄌ서시프
	프레시지	간편식	밀키트	784 <u>~</u>	쉬운요리	레시피박스	즉석식품
프레시지	1.000000	0.328985	0.451127	0.571446	0.506548	0.576270	0.444891
간편식	0.328985	1.000000	0.225857	0.303071	0.429705	0.110700	0.507771
밀키트	0.451127	0.225857	1.000000	-0.069997	0.030594	0.137312	0.052892
쿠킹박스	0.571446	0.303071	-0.069997	1.000000	0.652141	0.380749	0.569538
쉬운요리	0.506548	0.429705	0.030594	0.652141	1.000000	0.351722	0.669594
레시피박스	0.576270	0.110700	0.137312	0.380749	0.351722	1.000000	0.164078
즉석식품	0.444891	0.507771	0.052892	0.569538	0.669594	0.164078	1.000000

```
ax=plt.figure(figsize=(11,7))
ax=sns.heatmap(data=result12.corr(), annot=True, cmap='YIGn')
ax.xaxis.tick_top()
ax.xaxis.set_label_position('top')
ax.yaxis.tick_left()
ax.set_yticklabels(ax.get_yticklabels(),rotation=90)
plt.xticks(va='center',fontsize='10')
plt.yticks(va='center',fontsize='10')
plt.show()
```

5) 시계열 분석

데이터 랩에서 수집한 키워드 클릭량과 검색량 데이터로 시계열 그래프 생성

```
m = Prophet(yearly_seasonality=True, weekly_seasonality=True,
   daily_seasonality=True.)
m.fit(mealkit);
future = m.make_future_dataframe(periods=90)
future.tail()
452 2021-03-18
453 2021-03-19
                                   plt.figure(figsize=(12, 6))
454 2021-03-20
                                   plt.plot(mealkit_forecast, color='orange', Is = '--')
455 2021-03-21
                                   plt.plot(mealkit, color='orange')
                                   plt.plot(fresheasy_forecast, color='vellowgreen', ls = '--')
 456 2021-03-22
                                   plt.plot(fresheasy, color='vellowgreen')
                                   plt.grid()
                                   plt.legend(['밀키트 예측값', '밀키트 실제값',
                                               '프레시지 예측값', '프레시지 실제값'])
```



5. 분석결과 및 시각화



인지도 분석 결과

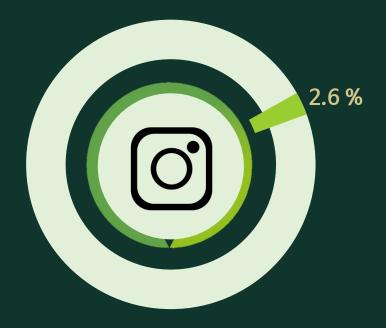
프레시지를 포함한 비율

전체 3600건 중 150건

4.2 %

프레시지를 포함한 비율

전체 2900건 중 75건





인지도 분석 결과

1. 인지도 분석에 사용한 방법이 프레시지의 경우와 맞지 않음

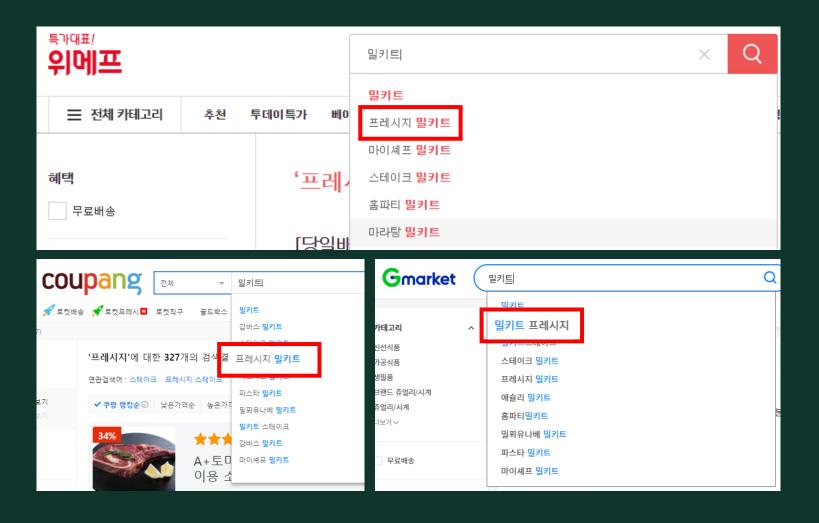
2. 프레시지는 인지도로 고객을 유입시키는 것이 아님

3. 프레시지는 언론 홍보비에 투자를 하지 않음



인지도 분석 결과

국내 오픈마켓, 소셜커머스 밀키트 연관검색어 최상단에 위치





인스타그램 프레시지 워드클라우드 분석결과

집밥, 요리스타그램, 홈쿡 → 밀키트 조리도 집밥의 범주에 들어감 감바스알아히오, 스테이크, 밀푀유나베 → 소비자가 많이 찾는 제품 혼밥, 신혼밥상, 자취밥상 → 주로 1~2인 가구에서 소비함 Jmt, 존맛, 맛집, 추천 → 프레시지 밀키트에 대한 긍정적인 반응





블로그+카페 프레시지 워드클라우드 분석결과



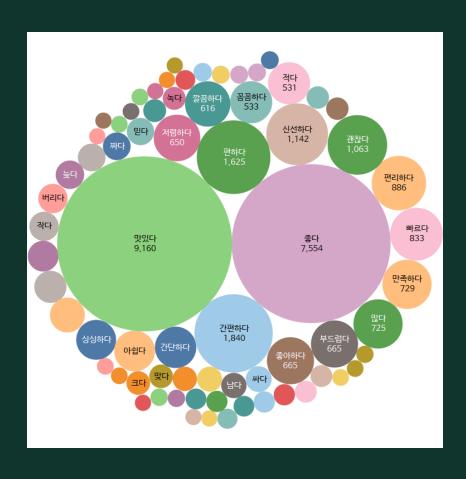
간편, 후기, 소스, 솔직,건강 → 상대적으로 다양한 키워드 국물, 매콤, 탄탄, 무배 → 자세한 표현





프레시지 태블로 리뷰 분석결과

가장 많이 등장한 키워드 → 좋다, 맛있다





프레시지 연관 검색어 상관관계 분석결과

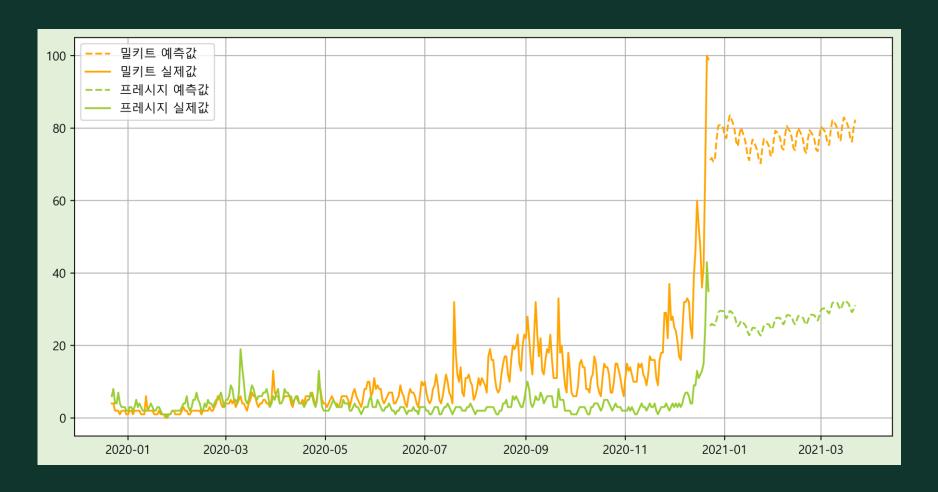
프레시지와 가장 높은 상관관계를 보이는 키워드 : 레시피박스(0.58) 가장 높은 상관관계를 보인 키워드 : 쉬운요리 - 즉석식품(0.67)

	프레신지	간편식	밀키트	쿠킹 <u></u> 박스	쉬운,요리	레시핀박스	즉석 _l 식품	- 1.0
프레시지	1	0.33	0.45	0.57	0.51	0.58	0.44	1.0
간편식	0.33	1	0.23	0.3	0.43	0.11	0.51	- 0.8
97月	0.45	0.23	1	-0.07	0.031	0.14	0.053	- 0.6
쿠킹박스	0.57	0.3	-0.07	1	0.65	0.38	0.57	- 0.4
쉬운요리	0.51	0.43	0.031	0.65	1	0.35	0.67	
레시피박스	0.58	0.11	0.14	0.38	0.35	1	0.16	- 0.2
즉석식품	0.44	0.51	0.053	0.57	0.67	0.16	1	- 0.0



프레시지 연관 검색어 클릭량 시계열 예측결과

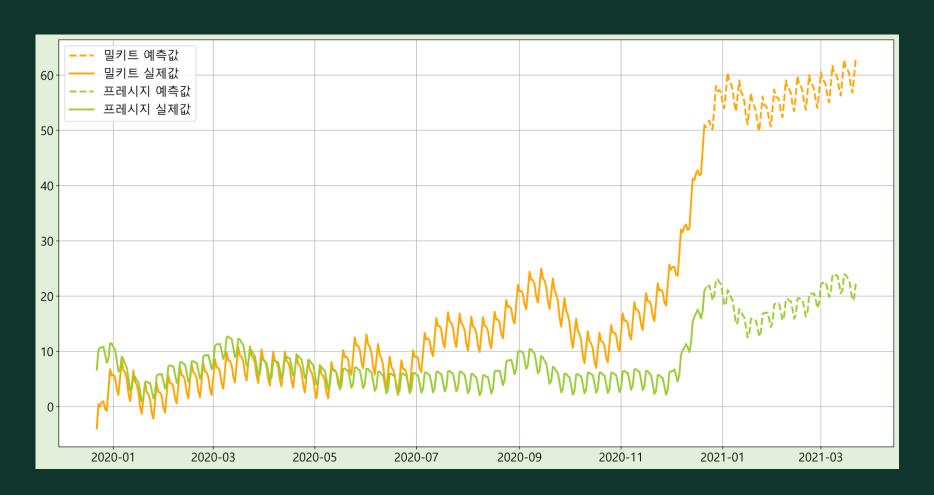
분석결과 향후 3개월 간 밀키트와 프레시지의 쇼핑 클릭량은 증가할 것으로 보임





프레시지 연관 검색어 검색량 시계열 예측결과

분석결과 향후 3개월 간 밀키트와 프레시지의 쇼핑 검색량은 증가할 것으로 보임



감사합니다.

참고자료

한국농촌경제연구원:

https://www.krei.re.kr/

금융감독원 전자공시 시스템:

http://dart.fss.or.kr/

프레시지 홈페이지:

https://www.fresheasy.co.kr/

통계청:

https://kosis.kr/

식품산업통계정보(가공식품 세분시장 현황):

https://www.atfis.or.kr/

한국농식품유통공사:

https://www.at.or.kr/

참고자료

네이버 데이터랩 쇼핑인사이트:

https://datalab.naver.com/shoppingInsight/sKeyword.naver

프레시지 관련 자료:

https://www.thebell.co.kr/free/Content/ArticleView.asp?key=202009171227235600109625&svccode=04

네이버 검색어 트랜드:

https://datalab.naver.com/keyword/trendSearch.naver