

품목과 가중치를 조정한 개인 맞춤형 체감물가지수 개발

스마트ICT융합학과 김대호
산업공학과 김태규
응용통계학과 박영원
산업공학과 송용재
경영학과 홍은기

CORPORATE
REPORT 7팀

목차

—

1. 소비자물가지수란?
2. 개인 맞춤형 체감물가지수
3. 시사점과 한계점

소비자물가지수란?

TABLE

— 소비자 물가지수



소비자물가지수란?

각 가정이 생활을 위해 구입하는 상품과 서비스의
가격 변동을 알아보기 위해 작성하는
통계 지수



품목

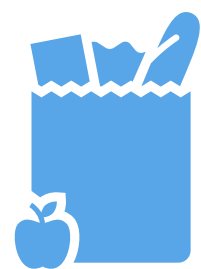


가격



가중치

— 소비자 물가지수



품목

458개



가격

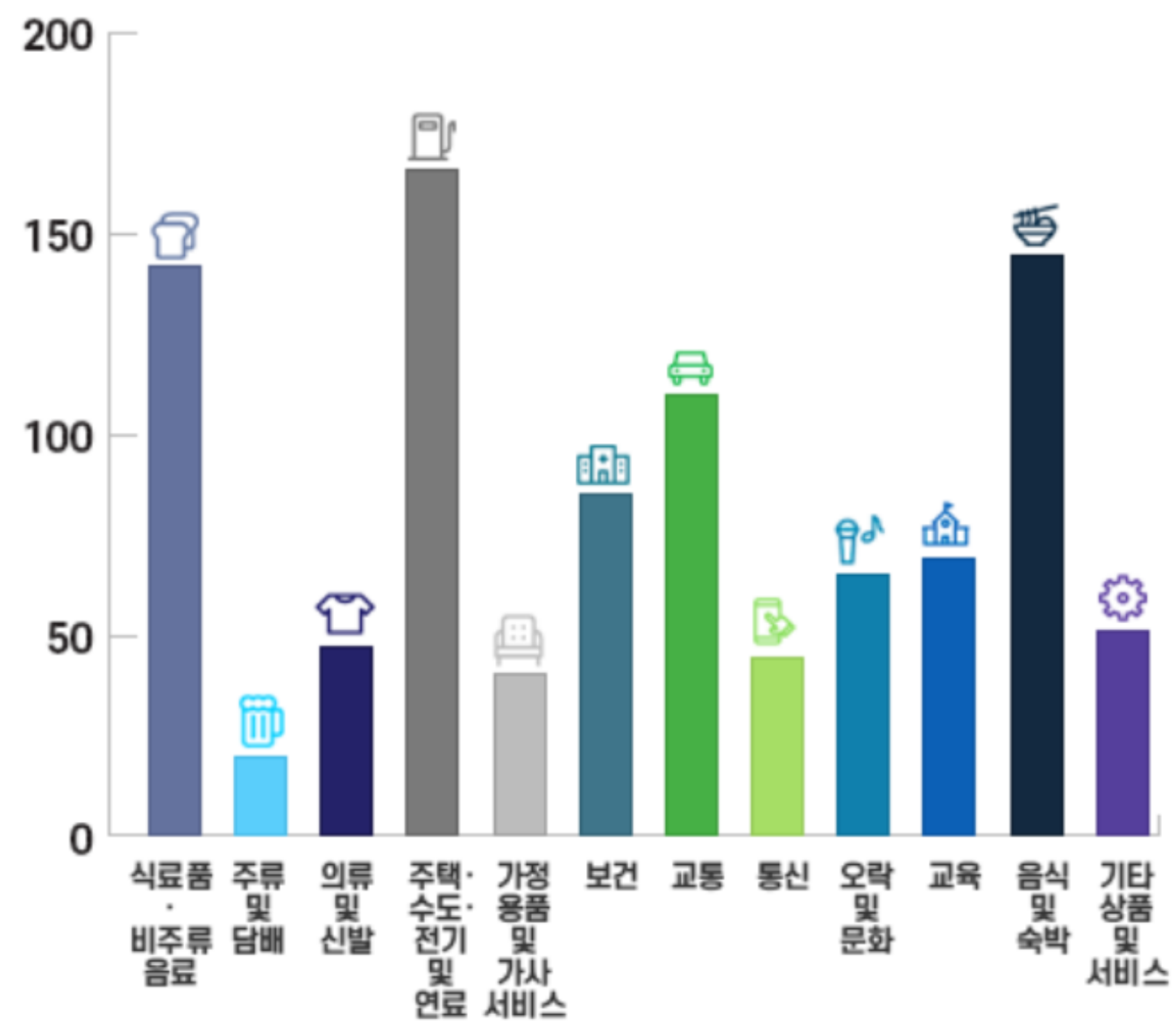
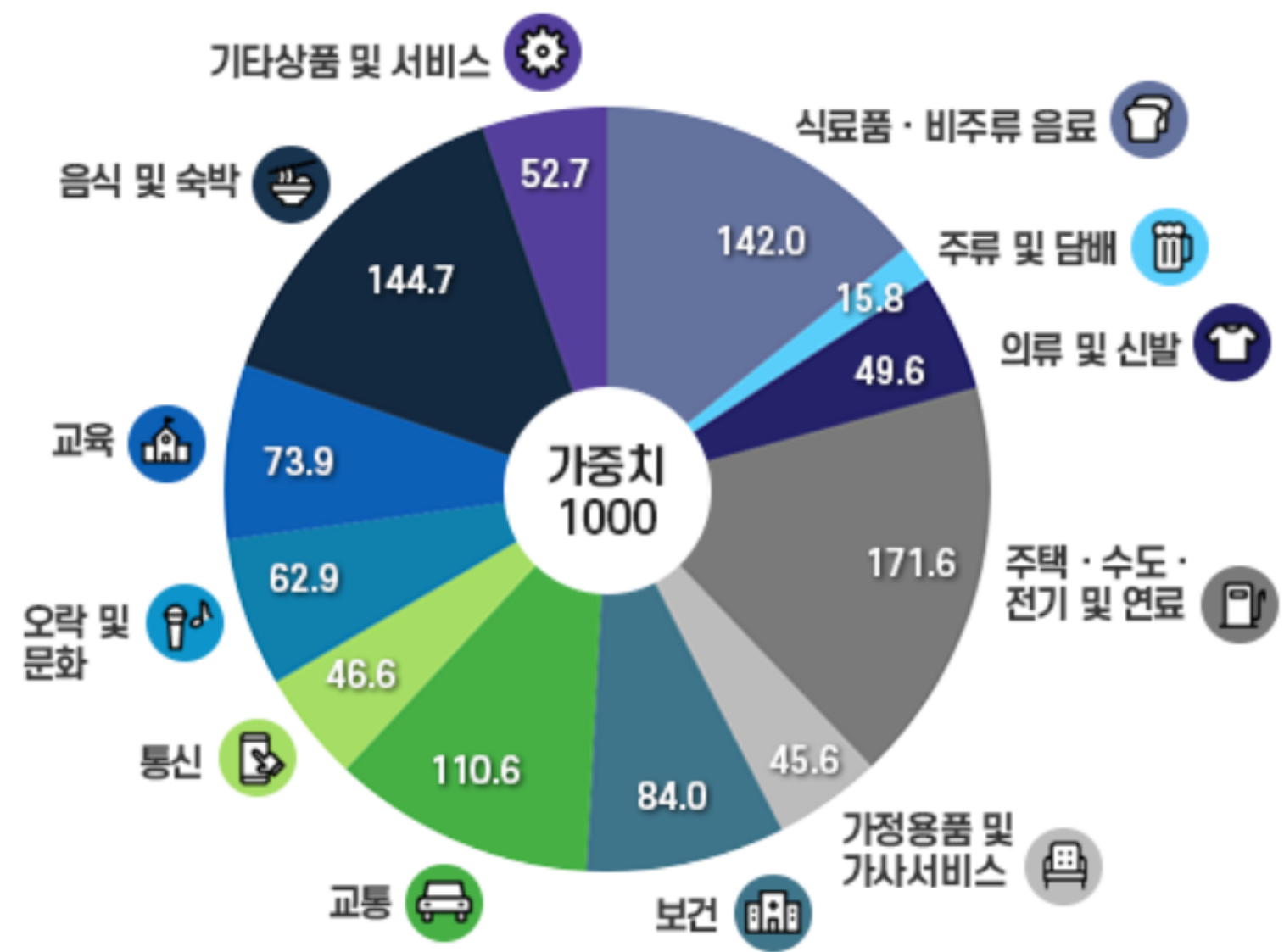
기준년도
대비 가격

(2020년)



가중치

해당 품목의
상대적 중요도



[2022년 기준 지출목적별 분류 가중치]

소비자 물가지수

연합뉴스 PICK · 4일 전 · 네이버뉴스

5월 소비자물가 2.7% 상승...2%대 둔화흐름 속 '金과일' 여전(중...
4일 통계청이 발표한 '5월 소비자물가동향'에 따르면 지난달 소비자물가
지수는 114.09(2
상승률은 올해 1



매일경제 PICK · 4일 전 · 네이버뉴스

[속보] 5월 소비자물가 2.7% 상승...두달 연속 2%대

4일 통계청이 발표한 '5월 소비자물가동향'에 따르면 지난달 소비자물가
지수는 114.09(2020년=100)로 지난해 같은 달보다 2.7% 올라다 소비자물



MT 머니투데이 · 3면 1단 · 4일 전 · 네이버뉴스

5월 소비자물가 상승률 2.7%...2개월 연속 2%대 물가

통계청이 4일 발표한 '2024년 5월 소비자물가동향'에 따르면 지난달 소비
자물가지수는 114.09(2020년=100)로 전년동월 대비 2.7% 상승했다. 지난
4월(2.9%)에 이어 2개월 연속 2%대 물가상승률을 기록했다. 품목별로 ...



소비자물가지수는 현재 국내에서 가장 공신력 있는 물가지수

=> 언론에서 물가를 다룰 때, 항상 인용되는 지표

2024 1/4분기
소비자물가지수

113.62

**‘소비자’ 물가지수는
‘소비자’가 실제 체감하는
물가를 잘 나타내고 있는가?**

—
VS
체감물가

1 너무 포괄적인 품목들

비흡연자



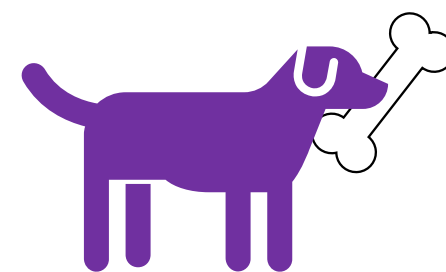
담배

독신 가구



학원비

비(比)반려인



반려동물
용품

—
VS
체감물가

이 곳에는 배경과
관련된 서브내용을
입력해 보세요.

2 가중치 갱신 주기가 너무 길다!



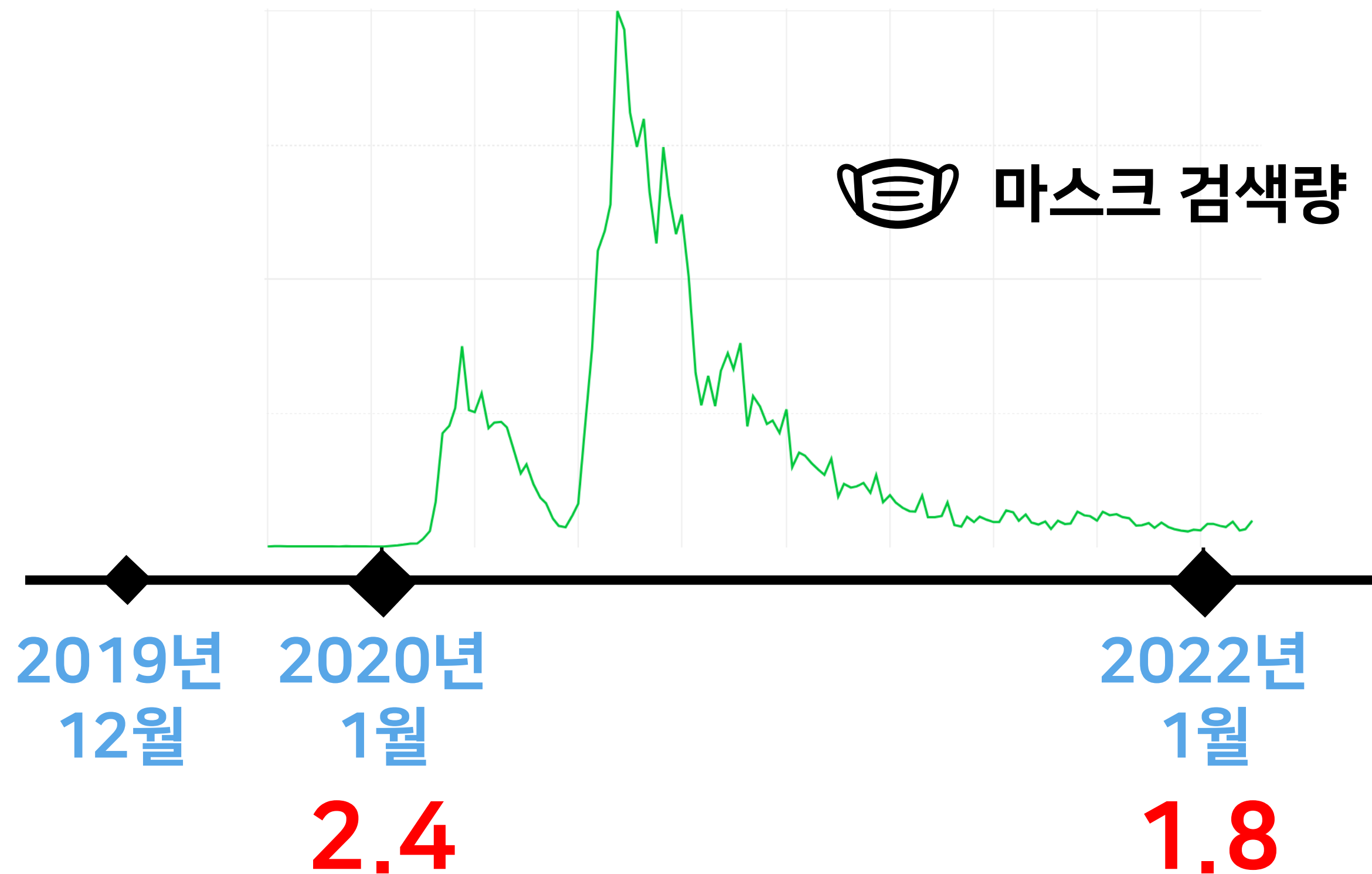
2020 채널A
뉴스특보

마스크 대란...
전국에서 이어진 핫걸음

BACKGROUND

—
VS
체감물가

2 가중치 갱신 주기가 너무 길다!



개인 맛춤형 체감물가지수

TABLE

프로젝트 목표

개인 맞춤형 체감물가지수

- 1 **개개인의 소비습관에 맞춰
필요 없는 품목 배제**
- 2 **물가지수 산출 시점의 품목
주목도를 반영하여 가중치 조정**

프로젝트 과정



설문조사
(소비습관 파악)

품목 선정



검색 데이터 수집



언론 데이터 수집

가중치 갱신



3명의 맞춤형
체감물가지수 산출

물가지수
산출

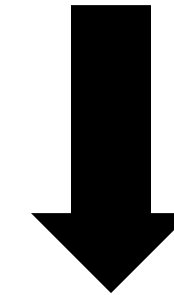
TABLE

설문지 작성

개인 맞춤형 소비자물가지수 품목 조사 설문지

1. 나는 남성이다. ☐ 네 ☐ 아니오
2. 나는 육류를 섭취하지 않는다. ☐ 네 ☐ 아니오
3. 나는 해산물을 섭취하지 않는다. ☐ 네 ☐ 아니오
4. 나는 패스트푸드(햄버거, 피자 등)에 관심이 없다. ☐ 네 ☐ 아니오
5. 나는 분식(떡볶이, 오뎅 등)에 관심이 없다. ☐ 네 ☐ 아니오
6. 나는 탕류(설렁탕, 갈비탕 등)에 관심이 없다. ☐ 네 ☐ 아니오
7. 나는 찌개류(김치/된장 찌개 등)에 관심이 없다. ☐ 네 ☐ 아니오
8. 나는 구내식당(학식)을 자주 이용하고 있다. ☐ 네 ☐ 아니오
9. 나는 편의점을 통해 식사를 해결하는 경우가 있다. ☐ 네 ☐ 아니오
10. 나는 술을 마시지 않는다. ☐ 네 ☐ 아니오
11. 나는 흡연을 하지 않는다. ☐ 네 ☐ 아니오
12. 나는 커피를 마시지 않는다. ☐ 네 ☐ 아니오

설문지를 토대로,
조사 대상자의 소비습관을 파악



총 458개 품목 중에,
대상자의 소비습관과 무관한
품목은 제외

—
설문지
작성

Ex)



20대 남성

자취(월세)

미성년 자녀 없음

술/담배 즐겨함

OTT 구독중

대학생

과일 구입 X

해외여행 X

반려동물 X

자차 없음

한의원 안 감

보험 가입

전세, 학원비, 승용차 유지비, 과일,
반려동물 용품/관리비, 항공료 등등 제외

—
가중치
조정

기간 : 2022/01 ~ 2024/04

검색 데이터

NAVER DataLab.

+

Google Trends



언론 데이터

KBS SBS

MBC

연합뉴스TV

YTN

...

검색 데이터

NAVER DataLab.



● 사과 | 과일

날짜	사과
2022-01-03	56.99397
2022-01-10	63.7379
2022-01-17	73.32803
2022-01-24	74.99649
2022-01-31	59.71397
2022-02-07	58.14833
2022-02-14	61.42449
2022-02-21	65.69145
2022-02-28	65.71481
2022-03-07	72.11758
2022-03-14	77.35196
2022-03-21	77.63705
2022-03-28	75.56199
2022-04-04	74.86096
2022-04-11	73.97298
2022-04-18	73.08501

■
■
■

사과
검색어

+ 비교

Google Trends



대한민국 ▼

22. 1. 1.~24. 4. 30. ▼

모든 카테고리 ▼

웹 검색 ▼

시간 흐름에 따른 관심도 변화 ②



사과	
2022-01-02	69
2022-01-09	71
2022-01-16	60
2022-01-23	59
2022-01-30	65
2022-02-06	76
2022-02-13	73
2022-02-20	61
2022-02-27	62
2022-03-06	62
2022-03-13	57
2022-03-20	61
2022-03-27	60
2022-04-03	63
2022-04-10	62
2022-04-17	63

— 동음이의어 처리

사과



사과
검색어

사과
과일



사과
행위

사과주
음료

언론 데이터

Crawling



URL을 활용하여
직접 크롤링

BeautifulSoup
,requests, urllib 등

크롤링 베이스 URL

```
base url = "https://search.naver.com/search.naver?where=news&query={query}&sm=tab_opt&sort=.&photo=2&field=0&pd=3&ds={start_date}&de={end_date}&docid=&related=0&mynews=0&office_type=0&office_section_code=0&news_office_checked=&nso=so%3Ar%2Cp%3Afrom{start_date}to{end_date}&is_sug_officeid=0&office_category=0&service_area=0"
```

&query = {query} : 검색어 지정

&photo = 2 : 동영상도 포함된 뉴스만 가져오기

&ds={start_date}&de={end_date} : 날짜

총 28개월을 4개월씩 7분할하여 크롤링 진행

```
# 각 키워드에 대해 두 날짜 범위를 사용하여 기사 수집하고 하나의 엑셀 파일에 저장
for keyword in keywords:
    print(f"Fetching articles for keyword '{keyword}'...")
    df1 = fetch_news_articles(base_url, keyword, "2024.01.02", "2024.04.30")
    time.sleep(5)
    df2 = fetch_news_articles(base_url, keyword, "2023.09.01", "2023.12.31")
    time.sleep(5)
    df3 = fetch_news_articles(base_url, keyword, "2023.05.01", "2023.08.31")
    time.sleep(5)
    df4 = fetch_news_articles(base_url, keyword, "2023.01.01", "2023.04.30")
    time.sleep(5)
    df5 = fetch_news_articles(base_url, keyword, "2022.09.01", "2022.12.31")
    time.sleep(5)
    df6 = fetch_news_articles(base_url, keyword, "2022.05.01", "2022.08.31")
    time.sleep(5)
    df7 = fetch_news_articles(base_url, keyword, "2022.01.01", "2022.04.30")
    time.sleep(5)
```

—
동음이의어
처리

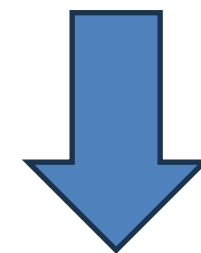
Chat-GPT 활용



facebook/bart



키워드 기반 텍스트 분류

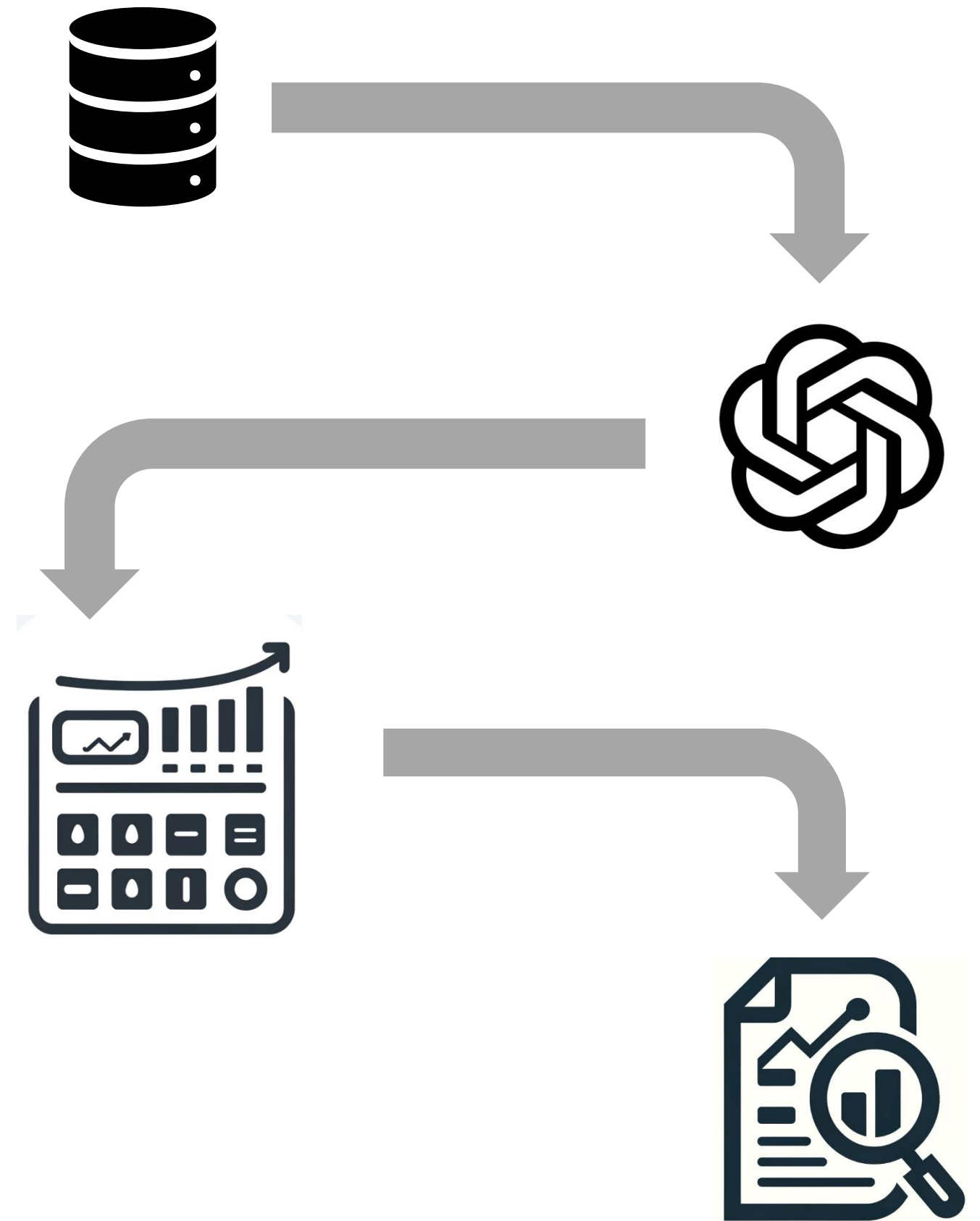


번역해서 분류하기

첫 번째 시도

Chat-GPT 활용

ChatGPT



두 번째 시도



facebook/bart

사과 문장 사과 의미

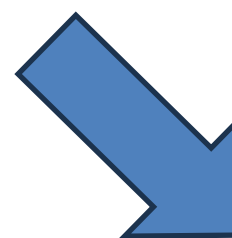
0 농촌 고령화로 오래된 과수원의 폐원이 늘면서 사과와 배 재배면적이 줄고 있습니다 통... apology
1 고민정·김의겸 등 문재인 정부 청와대 출신 더불어민주당 의원들은 29일 윤석열 대통... apology
2 숙명여대는 오늘(29일) '개인정보 유출에 대한 안내 및 사과문'이라는 제목의 게시... apology
3 민주당은 귀를 의심할 충격적 발언이라며 사과를 요구했지만, 국민의힘은 북한 선의에... apology
4 문재인 정부 청와대 출신 의원들은 긴급기자회견을 열고, 대통령의 편협한 사고 체계... apology
5 이를 전 채널A가 보도한 곰팡이 투성이 임대주택과 관련해 SH 서울주택도시공사가 공... apology
6 이 대표는 "박 위원장은 당장 망언에 대해 국민께 사과하고 물러나야 한다 apology
7 민주당 의원들, 논란 커지자 사과\n... apology
8 [기자]< 불심에 '화들짝' >중국 초청으로 티베트를 방문했던 더불어민주당 의원들이... apology
9 이주호 교육부 장관은 대통령을 '입시 전문가'로 치켜세운 발언과 관련해 오해가 있었... apology
11 그런 부분에서 오해가 있었다면 사과드리겠습니다 apology

최종 결과

키워드 기반 텍스트 분류

```
stop_keywords = [  
    '사과하고', '사과하며', '사과해서', '사과하니', '사과한다', '사과했다', '사과해도', '사과하는', '사  
    '사과한대도', '사과할까', '사과했지만', '사과해야', '사과해라', '사과하는데', '사과한지', '사과하리  
    '사과해서라도', '사과한다면', '사과했을까', '사과하자', '사과했으니', '사과하는구나', '사과했는데도  
    '사과합시다', '사과해볼까', '사과해봤자', '사과하는가', '사과했으면', '사과하는데도', '사과하자마자  
    '사과했겠지', '사과했었다', '사과할게요', '사과하느냐', '사과해버리다', '사과하더니', '사과하려면',  
    '사과할텐데', '사과할래요', '사과했겠지만', '사과하는데서', '사과문', '사과드리겠습니다', '사과부터  
    '사과를 요구', '사과를 촉구', '사과드립니다', '송구', '송구하다', '송구합니다', '죄송', '죄송하단'  
    '사과 없이', '사과는 없었습니다', '사과를 받고', '사과했습니다', '거듭', '사과 마땅', '사과를 하겠
```

번역할 데이터 양을
4분의 1로 줄임



번역 후 분류

URL	Title	Content	Date	Title_English	Content_English	Apple_Count	Apology_Count
https://n.n	사과·배, 농촌 고령화	농촌 고령화로 오래	2023.06.30. 오전 10	Apples, pears, rural	As rural aging has i	2	0
https://n.n	[기후는 말한다] "더 [앵커] 수산물뿐 아	니	2023.06.21. 오후 12	[The climate says] "	[Anchor] The crops	6	0
https://n.n	"더 올라갈 곳도 없 [앵커] 바다뿐 아니	리	2023.06.20. 오후 9:	"There is no place t	[Anchor] The crops	6	0
https://n.n	[지구촌 더뉴스] 11 [앵커] 세계 최대 IT		2023.06.22. 오전 10	[Global Village The	[Anchor] Apple, the	10	1
https://n.n	"수정률 절반 감소". [KBS 청주] [앵커] 총		2023.05.05. 오전 8:	"Reduction of half t	[KBS Cheongju] [An	9	0
https://n.n	"수정률 절반 감소". [KBS 청주] [앵커] 총		2023.05.04. 오후 9:	"Reduction of half t	[KBS Cheongju] [An	9	0
https://n.n	널뛰는 날씨에 '사과 [앵커] 요즘 큰 일교차		2023.05.01. 오후 7:	The apple tree was	[Anchor] It's easy to	5	0
https://n.n	휴전선까지 올라간 <앵커> 전국에서 사		2022.09.13. 오후 8:	'The land of apple'	<Anchor> The mos	7	2

점수
처리하기

1월 사과 언급

	기사 1	1회	$\sqrt{1}$
	기사 2	3회	$\Rightarrow \sqrt{3} \Rightarrow$
	기사 3	4회	$\sqrt{4}$

4.73

개별 기사에서 키워드가 언급된 횟수에
루트를 씌우고, 총합 계산!

—
모델
소개하기

PROPHET

1월 쌀 검색량

76

=

추세

+

주기성

+

오차항

64.06

9.01

2.93

BACKGROUND

PROPHET

—
모델
소개하기

추세

=

실제 값

-

주기성

-

오차항

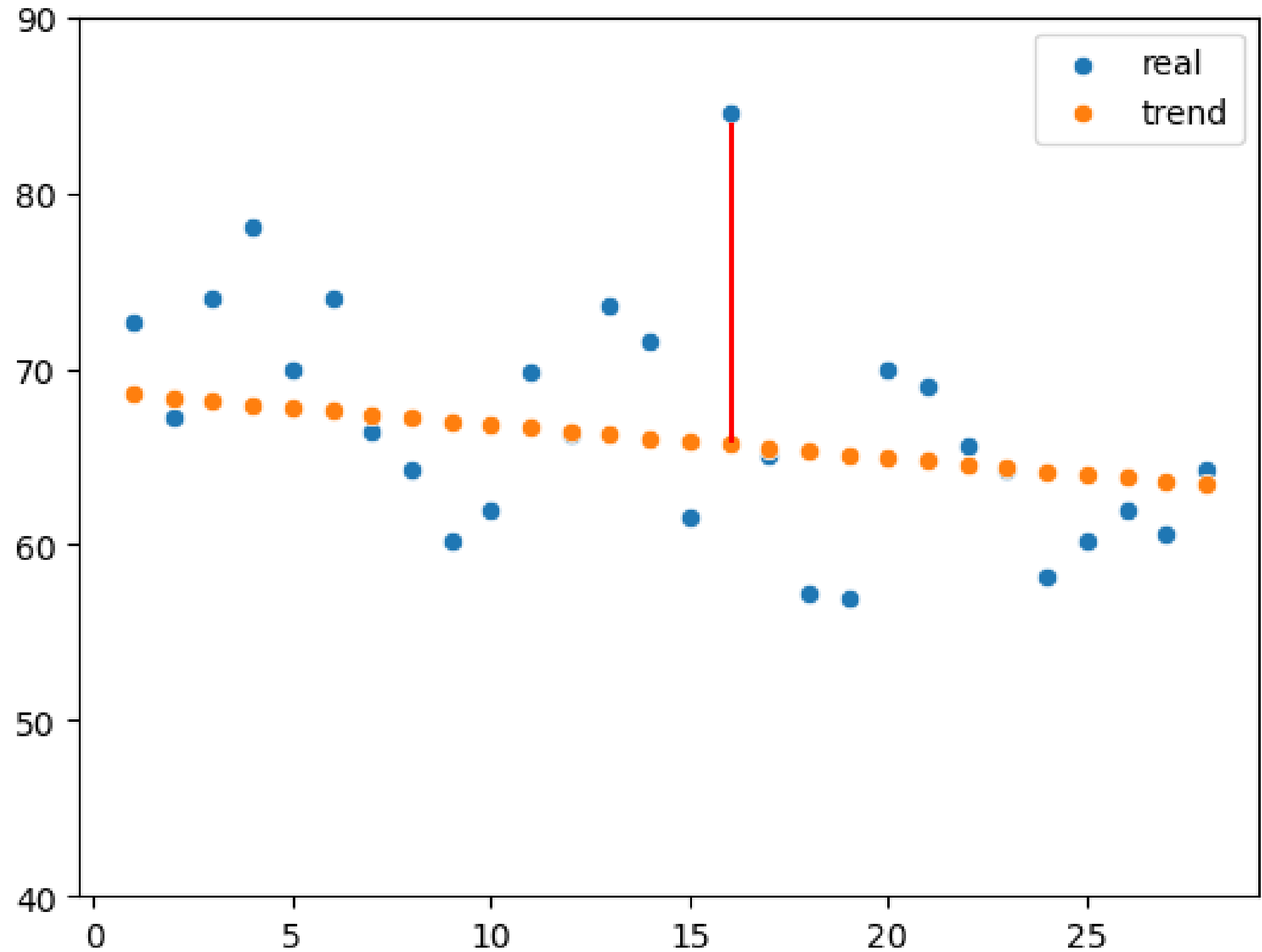
즉, 실제 값에서 계절성과 예측 불가능한 변수의 영향을 제거하면 '추세'가 된다!

—
방법론
소개하기

$$\frac{\text{실제 값}}{\text{추세}} = \frac{76}{64.06} = 1.19$$

계절성과 예측 불가능한 변수로 인해
추세 대비 얼마 만큼의 주목도를 보였는가

방법론 소개하기



▲ 28개월 간의 밀가루 검색량과 추세(구글)

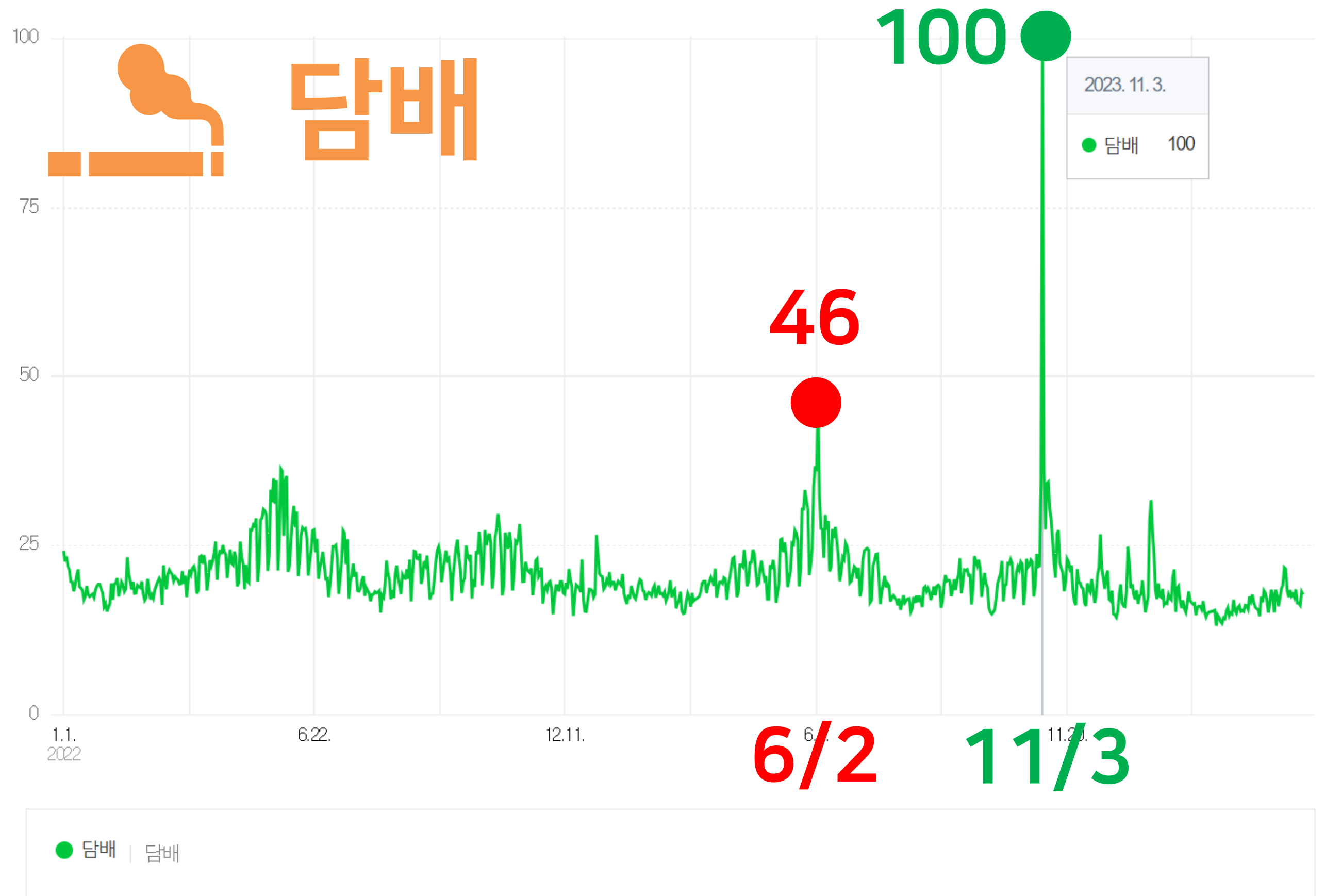
사례 소개하기

자막뉴스 메타 지구 **KBS NEWS**

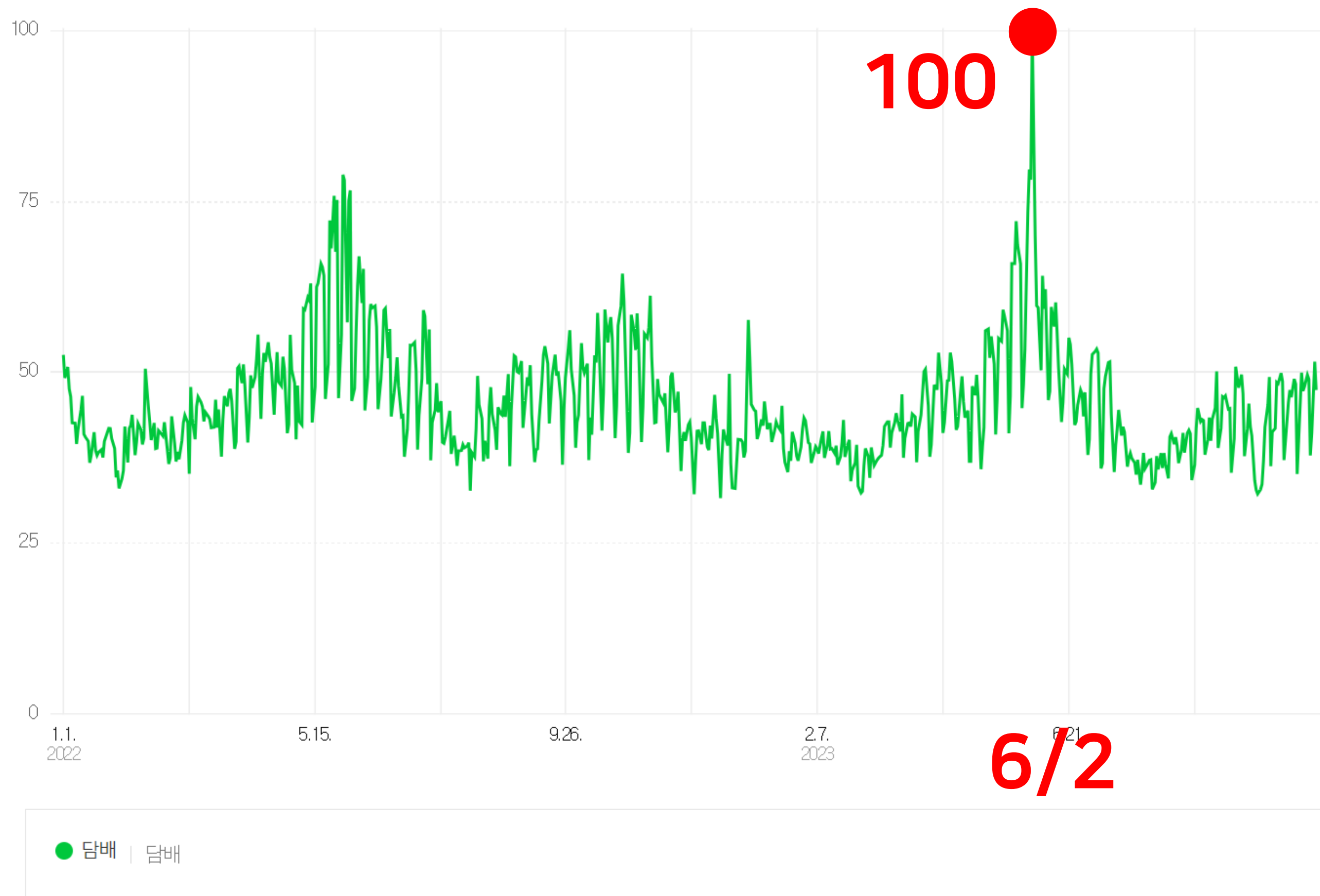
인간 영향 지구 **온난화 발생** ✓ 자연 상태 지구 **온난화 없음**

'메타 지구'와 비교해보니
"태풍 더 세졌다"

데이터 문제점



데이터 문제점



실제 값

추세

=

[1월 실제 검색량] / ~~[고점]~~

[1월 추세 검색량] / ~~[고점]~~

=

[1월 실제 검색량]

[1월 추세 검색량]

시점으로부터
독립!

[도시가스]의 가중치 구하기 (2024년 1월)



점유율

(1.25 x 0.16

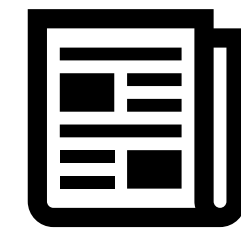
+



점유율

1.18 x 0.31

+



점유율

1.32 x 0.53)

도시가스
기본 가중치

x 11.5

=

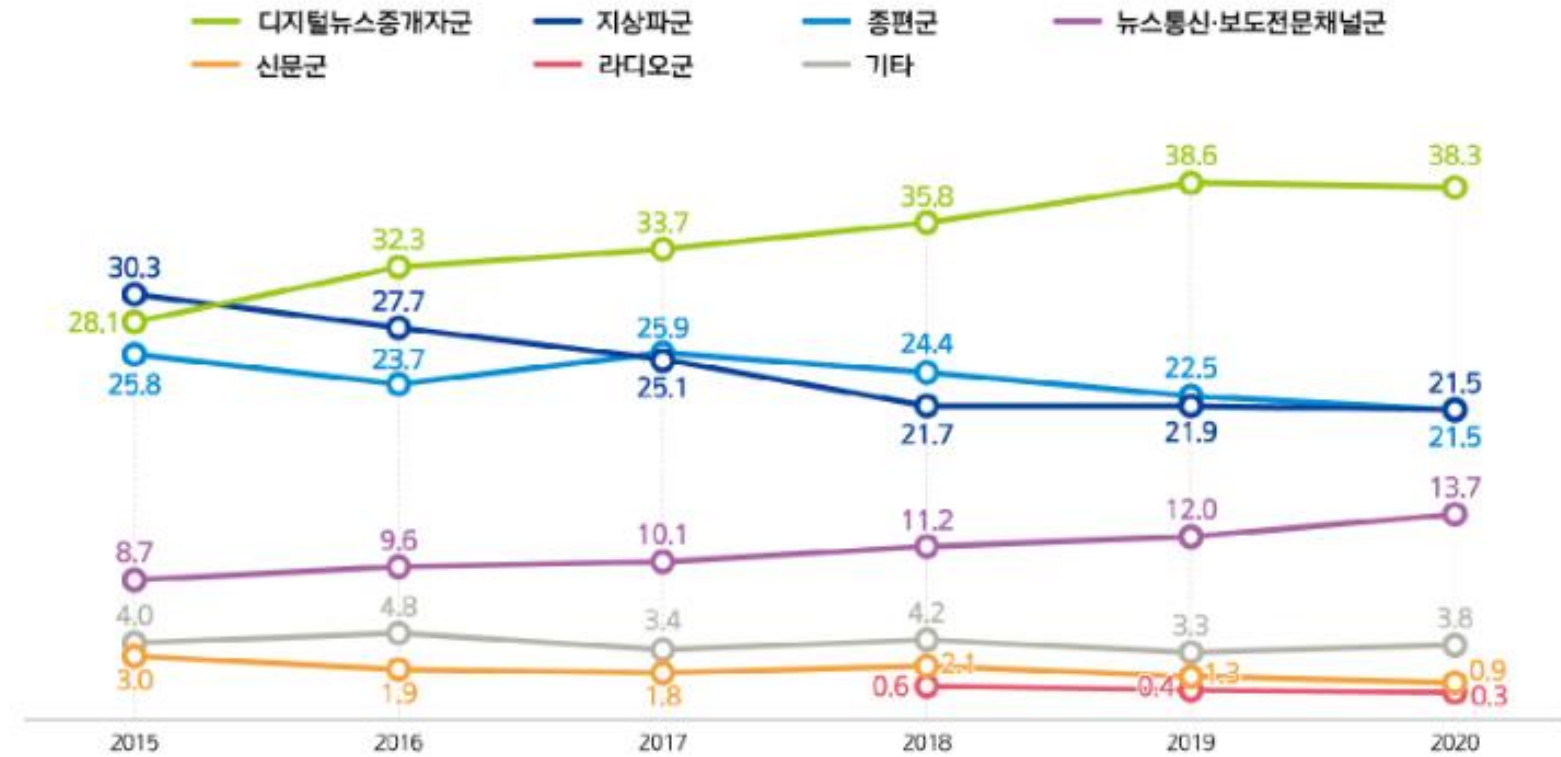
새로운 가중치

14.26



문화체육관광부

여론집중도조사 (2022년 발표)



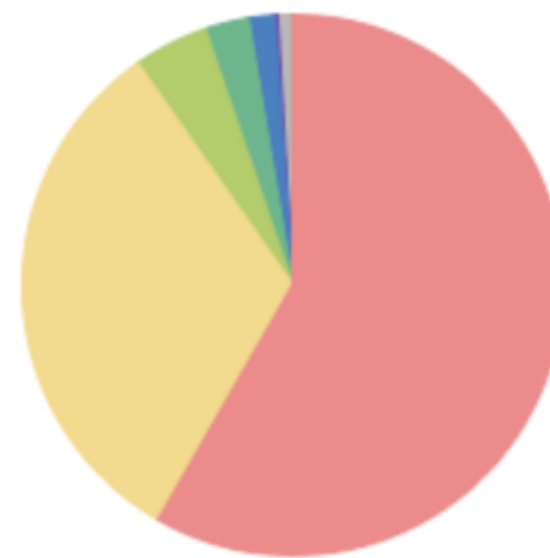
* '21년은 포털앱의 개인정보 보호 강화로 뉴스 이용량 측정이 어려움에 따라 조사 불가

INTERNET TREND™

2024년 1~4월 기준
포털 점유율

Total : 20,368,880





LOGGER™



- NAVER(네이버) (58.28 %)
- GOOGLE(구글) (32 %)
- DAUM(다음) (4.46 %)
- MSbing(마이크로소프트) (2.63 %)
- 기타 (1.53 %)
- ZUM(줌) (0.2 %)
- 나머지항목 (0.90 %)

BACKGROUND

가중치 갱신하기

품목(시점)	기존 가중치	갱신된 가중치
 라면 (1월)	3.1	→ 3.11
 비빔밥 (1월)	2.4	→ 1.98
 고등어 (4월)	1.85	→ 3.7
 수신료 (2월)	2.4	→ 5.62



품목별 소비자물가지수 (2024.01 ~ 2024.04)

1) 품목별 소비자물가지수(품목성질별: 2020=100) 「소비자물가」

수록기간 : 월, 분기, 년 1975.01 ~ 2024.05 / 자료갱신일: 2024-06-04 / [주석정보](#)

[시점](#) [증감률](#) [행렬전환](#) [열고정해제](#) [새 탭 열기](#) [화](#)

(단위: 2020=100)

시도별	품목별	2024.01	2024.02	2024.03	2024.04
<div>▲ ▼ ▢</div>	<div>▲ ▼ ▢</div>	<div>▲ ▼ ▢</div>	<div>▲ ▼ ▢</div>	<div>▲ ▼ ▢</div>	<div>▲ ▼ ▢</div>
전국	총지수	113.15	113.77	113.94	113.99
	농축수산물	122.71	125.98	126.54	123.54
	(농산물)곡물	104.76	103.79	102.58	100.20
	쌀	104.96	103.26	101.24	97.65
	현미	101.98	102.45	102.77	102.67
	잡쌀	99.60	99.46	100.75	100.62
	보리쌀	118.80	118.72	120.00	121.66
	콩	101.76	103.88	103.37	103.21
	땅콩	113.96	114.36	115.76	116.61
	혼식곡	108.58	108.74	109.18	109.41
	(농산물)채소	123.65	132.00	131.90	124.29
	배추	78.97	83.49	93.96	117.95
	상추	115.82	112.38	109.99	92.85
	시금치	101.82	125.41	104.23	94.70
	양배추	88.05	86.18	91.56	127.94
	미나리	176.36	176.07	148.23	121.77
	깻잎	147.24	145.11	136.46	110.57
	부추	179.16	205.90	203.38	143.59
	무	82.77	82.54	85.33	90.69
	열무	103.19	110.61	136.11	117.75
	당근	84.12	90.20	109.77	139.13
	감자	126.38	134.73	146.36	159.43
	고구마	95.08	92.78	94.74	97.77
	두릅	122.81	122.51	124.81	124.22

소비자물가지수 계산법

전국 물가지수의 작성

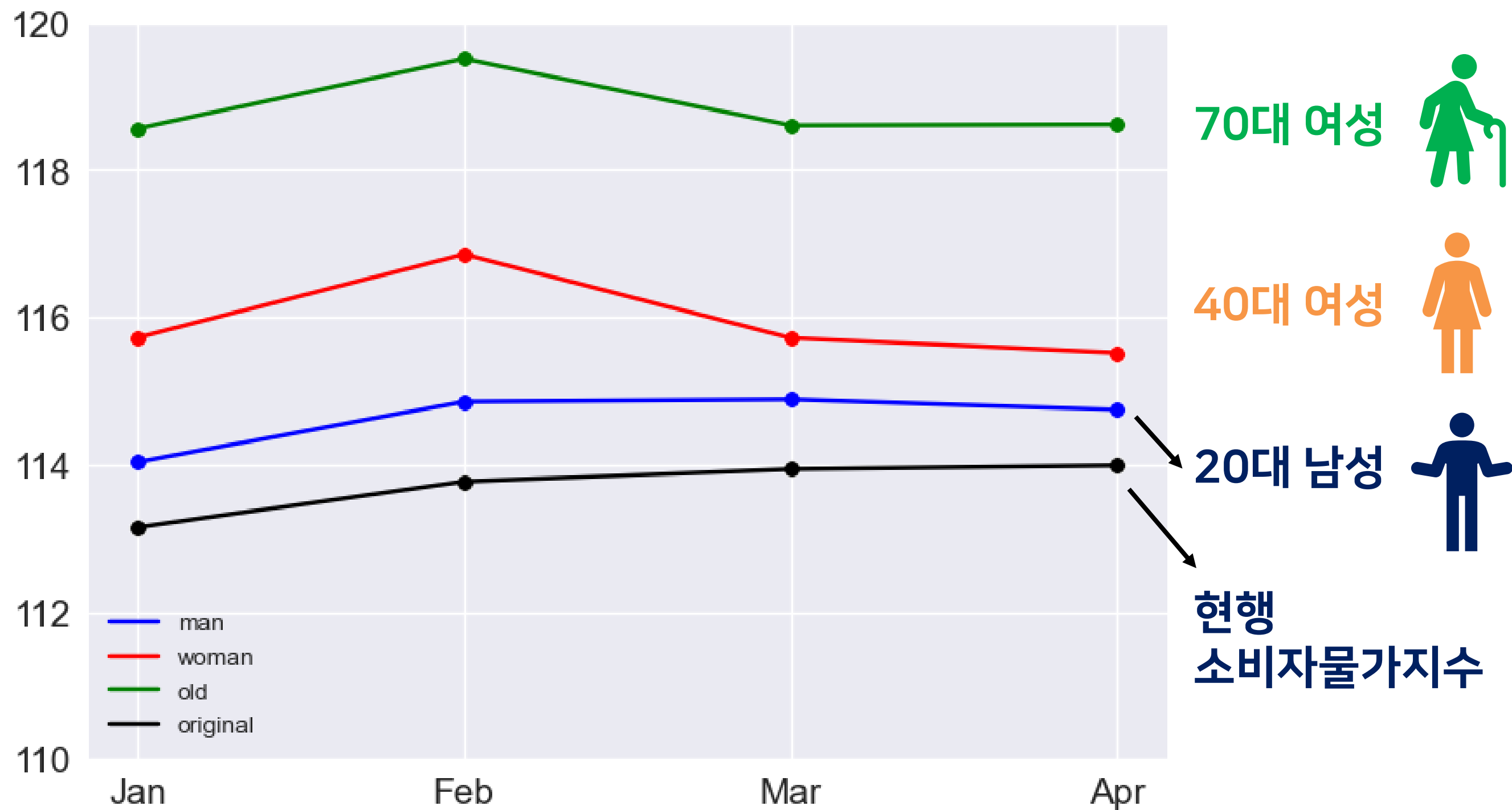
소비자물가지수 산출 예

	9월	가중치
쌀(20Kg)	124.76	400
햄버거(1개)	102.85	200
휘발유(1L)	109.35	400
계	114.21	1,000

$$9\text{월 소비자물가 지수} = \frac{124.76 \times 400 + 102.85 \times 200 + 109.35 \times 400}{1,000} = 114.21$$

지수 산출하기

2024년 개인 맞춤형 물가지수



20대 남성 

담배

대중교통

PC방

40대 여성 

과일

자녀 교육비

항공료

70대 여성 

치과(보철)

수산물

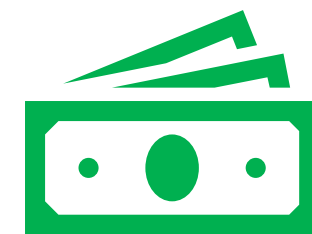
화초

프로젝트의 시사점과 한계점

TABLE

프로젝트 시사점

- 1 **현행 소비자물가지수의 한계점을 인식하고, 이를
민간차원에서 개선하려고 했던 시도는 처음**
- 2 **같은 인플레이션 상황에서도,
특정 계층에게 더 큰 영향이 갈 수 있음을 보임**
- 3 **향후 금융업계의 마이데이터 서비스와
연계하면, 전 국민으로 대상으로 개인 맞춤형
체감물가지수 제공 가능**



결제내역/
소비습관
데이터 활용



품목 선택, 가중치 조정



○○○ 님
121.76

프로젝트 한계점

1

원래부터 소비자물가지수 산출에 누락된 품목들을
새로이 추가하지 못했음

2

개개인마다 각 품목별로 느끼는 중요성의 차이도
모두 다른데, 이를 가중치에 담아내지 못함

Q & A

TABLE

부록

T A B L E

부록1 네이버 구글 검색량 수집 데이터

A	B	C	D	E	F	G	H	I	J	K	L	M
날짜	쌀	밀가루	라면	시리얼	빵	닭	햄	갈치	명태	굴비	고등어	우유
01/02/2022	17.79439	9.314223	8.308764	36.17792	41.09654	9.377834	41.87257	8.118409	44.0423	5.170724	25.10939	44.5086
01/09/2022	18.96285	10.26173	8.825661	43.889	43.46474	9.91489	48.30083	11.24356	45.44625	7.678397	30.66115	51.0798
01/16/2022	19.12075	10.0424	8.345083	43.1151	44.31154	10.94769	48.14838	11.36144	44.16993	8.690847	30.71175	54.93706
01/23/2022	17.60325	10.29682	7.860821	41.13504	45.08659	10.16276	51.48319	12.59322	41.22527	10.06098	30.63138	50.48504
01/30/2022	12.16487	8.855091	6.550683	31.71581	39.76652	7.791451	39.92885	6.725443	51.32646	7.871894	19.44452	41.86973
02/06/2022	19.98504	9.267431	8.530367	44.02844	44.57946	9.245634	46.11573	7.027374	41.43495	5.806927	27.20804	56.69823
02/13/2022	18.52073	9.600813	8.467729	48.00948	45.45498	10.52631	45.06129	7.156084	39.61163	4.985043	27.47	57.1544

2024-03-17	80	68	75	59	33	59	38
2024-03-24	68	57	76	59	33	55	39
2024-03-31	65	54	75	64	30	70	39
2024-04-07	60	67	76	61	34	47	41
2024-04-14	81	70	71	63	29	64	40
2024-04-21	58	60	75	60	32	51	36
2024-04-28	66	60	74	60	37	56	42

부록2 품목별 기사 크롤링

 ott_results.xlsx	 마늘_results.xlsx	 스파게티_resu
 pc방_results.xlsx	 마스크_results.xlsx	 승용차_results
 tv_results_unique.xlsx	 막걸리_results.xlsx	 시계_results.xl
 가구_results.xlsx	 맥주_results.xlsx	 시금치_results
 가방_results.xlsx	 명태_results.xlsx	 시리얼_results
 간장_results.xlsx	 문제집_results.xlsx	 식용유_results
 갈치_results.xlsx	 미역_results.xlsx	 식초_results.xl
 감자_results.xlsx	 밀가루_results.xlsx	 신문_results.xl
 계란_results.xlsx	 바나나_results.xlsx	 쌀_results.xlsx
 고구마_results.xlsx	 바디워시_results.xlsx	 쌀국수_results
 고등어_results.xlsx	 반려동물_results.xlsx	 쓰레기봉투_re
 고추_results.xlsx	 밥술_results.xlsx	 약국_results.xl
 고추장_results.xlsx	 배추_results.xlsx	 양념_results.xl
 고춧가루_results.xlsx	 버섯_results.xlsx	 양파_results.xl
 과자_results.xlsx	 버스_results.xlsx	 에어컨_results
 관람_results.xlsx	 병원비_results.xlsx	 여행_results.xl
 관리비_results.xlsx	 보청기_results.xlsx	 열무_results.xl
 굴비_results.xlsx	 보험_results.xlsx	 염색약_results
 굴_results.xlsx	 복숭아_results.xlsx	 영양제_results
 기숙사_results.xlsx	 볶음밥_results.xlsx	 영화_results.xl
 김밥_results.xlsx	 부동산_results.xlsx	 오렌지_results
 김치_results.xlsx	 부엌_results.xlsx	 오이_results.xl
 김치찌개_results.xlsx	 불고기_results.xlsx	 요양_results.xl
 난방비_results.xlsx	 브로콜리_results.xlsx	 우산_results.xl
 냉면_results.xlsx	 비누_results.xlsx	 우유_results.xl
 냉장고_results.xlsx	 비빔밥_results.xlsx	 운동화_results
 노래방_results.xlsx	 빵_results.xlsx	 월세_results.xl

	A	B	C	D	E
1	URL	Title	Content	Date	
2	https://	내일부터 '내일(1일)부터 병	2024.04.30. 오후 11:19		
3	https://	인천시립도 ■ 제보하기▷ 전	2024.04.30. 오후 9:46		
4	https://	청양군, 3면[KBS 대전] [앵커]	2024.04.30. 오후 9:40		
5	https://	청양군, 3면[KBS 대전] [앵커]	2024.04.30. 오후 9:40		
6	https://	"일할 사람[앵커] 청양			
7	https://	"30대가 시[앵커] 청양			
8	https://	[인력절벽][앵커] 청양			
9	https://	"30대가 시[앵커]저출			
10	https://	"20년 전까[KBS 부산			
11	https://	"20년 전까[KBS 부산			
12	https://	강제 노역[KBS 부산			
13	https://	[문화현장] [FunFun			
14	https://	[더 보다] [더보다 9			
15	https://	병원도 마[앵커]이저			
16	https://	코로나19 [앵커] 정			
17	https://	옥천군립 [KBS 청주			
18	https://	[더 보다] [더보다 8			
19	https://	백세시대리[시사기획			
20	https://	인구 5천 뒤[앵커]저출			

	A	B	C	D	E
1	URL	Title	Content	Date	
2	https://	[와글와글] [뉴스투데이]인터넷	2023.06.30. 오전 6:34		
3	https://	다들단브타다들단브타 지하철	2023.06.28. 오후 4:53		
4	https://	다들단브타다들단브타 지하철	2023.06.30. 오전 9:12		
5	https://	다들단브타다들단브타 지하철	2023.06.29. 오전 11:1		
6	https://	다들단브타다들단브타 지하철	2023.06.29. 오전 11:0		
7	https://	다들단브타다들단브타 지하철	2023.06.29. 오전 6:07		
8	https://	다들단브타다들단브타 지하철	2023.06.29. 오전 7:24		
9	https://	다들단브타다들단브타 지하철	2023.06.28. 오후 8:44		
10	https://	다들단브타다들단브타 지하철	2023.06.28. 오후 9:22		
11	https://	다들단브타다들단브타 지하철	2023.06.27. 오후 7:47		
12	https://	다들단브타다들단브타 지하철	2023.06.28. 오전 8:05		
13	https://	다들단브타다들단브타 지하철	2023.06.30. 오전 6:37		
14	https://	다들단브타다들단브타 지하철	2023.06.29. 오전 6:20		
15	https://	다들단브타다들단브타 지하철	2023.06.29. 오후 3:55		
16	https://	다들단브타다들단브타 지하철	2023.06.27. 오후 7:06		
17	https://	다들단브타다들단브타 지하철	2023.06.22. 오후 10:5		

	A	B	C	D	E
1	URL	Title	Content	Date	
2	https://	'우크라 사태' 긴	미국이 러시아에	2022.02.25. 오후 7:24	
3	https://	고흥 관내 고교	고흥 지역 고등학	2022.02.23. 오후 6:38	
4	https://	울산시, 코로나19	[KBS 울산]울산시	2022.02.25. 오후 11:10	
5	https://	[World Now_영광	원숭이 한 마리가	2022.02.14. 오후 12:03	
6	https://	정부 "대러 전략<앵커>우리 정부		2022.02.28. 오후 5:10	
7	https://	[자막뉴스] 바이	조 바이든 미국	2022.02.25. 오후 1:48	
8	https://	"딸들을 전쟁터	김현기 순회특파	2022.02.28. 오전 5:00	
9	https://	충남북부 1월 수	[KBS 대전]천안,	2022.02.21. 오전 8:31	
10	https://	분뇨·하수·음식쓰	<앵커>음식물쓰	2022.02.28. 오후 9:19	
11	https://	우크라이나는 '하	[뉴스데스크] ◀	2022.02.19. 오후 8:24	
12	https://	[이슈톡] 80년간	[뉴스투데이]미국	2022.02.08. 오전 7:28	
13	https://	'안보불안' 우려	F-35A·조음속순항	2022.02.28. 오후 5:47	
14	https://	울산시 정보보안	[KBS 울산]울산시	2022.02.07. 오전 7:48	
15	https://	"러시아 전략물자	[뉴스데스크] ◀	2022.02.28. 오후 8:20	
16	https://	[친절한 경제] 우	<앵커>친절한 경	2022.02.28. 오전 9:40	
17	https://	정부, 한발 늦게	우리 정부는 오늘	2022.02.28. 오후 7:26	
18	https://	하드디스크 속	[사이언스샷]사기	2022.02.03. 오전 11:31	
19	https://	빌딩 주차장 '해	[앵커]요즘 대형	2022.02.27. 오후 6:44	
20	https://	기아, 미국서 41	기아가 에어백	2022.01.29. 오전 11:13	

부록3 동음이의어 처리 결과

6618	https://	44주년 5·18 행사취급외	[앵커]한애호 44주년이	2024.03.19.	오전 9:14		
6619	https://	[총선] 야 "박근혜도 권좌	[앵커]이재명 더불어민주당	2024.03.19.	오후 9:10		
6620	https://	이재명 "대책 없이 경기	[편상욱의 뉴스브리핑]인	2024.03.25.	오후 3:34		
6621	https://	'5·18 폄훼' 도태우 공천	[뉴스투데이] ◀ 앵커 ▶ 5·	2024.03.15.	오전 6:04		
6622	https://	[정치쇼] 공영운 "경쟁자	- 출마의 변? 경제국가대	2024.03.15.	오전 10:10		

1600	https://	[자막뉴스] 비트코인 32년 만에 최고 상승률	2024.03.12. 오전 8:16	사과는 지난달보다 7	1
1601	https://	[단독] 동료 시켜 '한[앵커]한 종합병원 팀장	2024.03.25. 오후 7:59	팀장에게 사과도 받지	1
1602	https://	아버지 레드향 섞이[KBS 제주] [앵커] 피땀	2024.03.25. 오후 7:11	[송종걸/레드향 판매	1
1603	https://	李 '전 국민 25만 원■ 방송 : 채널A 뉴스 T	2024.03.25. 오후 7:00	지금 보면 사과 가격	1
1604	https://	"벌 붕괴 현상" 위기<앵커>꿀벌이 집단으로	2024.03.25. 오후 5:31]전문가들은 꿀벌 실	1

397	https://	[이시각해]	■ 2월 취	2024.03.13	■ 사과 10	1	[Headline	■ The nu	2	0
398	https://	금사과 대	[앵커] 요	2024.03.16	[앵커] 요	1	I eat orang	[Anchor] M	2	0
399	https://	금사과 대	[앵커] 요	2024.03.15	[앵커] 요	1	I eat orang	[Anchor] A	1	0
400	https://	[직설] 과일	■ 용감한	2024.03.12	■ 용감한	1	[Direct] M	■ Brave T	6	0

부록4 검색량/기사 데이터 전처리

E	F	G	H	I	J	K	L	M	N
가구_trend	가방_trend	간장_trend	갈치_trend	감자_trend	계란_trend	고구마_trend	고등어_trend	고추_trend	고추장_trend
0.852633	0.8076	0.932935	0.957871	1.084272	0.851154	1.683148	1.551643	1.309049	1.02445
1.010693	0.935292	1.734119	1.935464	1.398343	1.106802	0.951132	1.463108	1.212477	0.53848
0.912839	1.103631	1.072528	1.831798	0.936317	1.223929	1.718248	1.222453	1.415165	1.009622
0.98669	1.080989	0.926216	2.241919	0.613188	1.705295	1.241018	2.592384	1.132796	1.049114

B	C	D	E	F	G	H	I
ott_trend	pc방_trend	tv_trend	가구_trend	가방_trend	간장_trend	갈치_trend	감자_trend
1.032739	0.966864	1.135818	1.133741	1.042228	1.056134	1.080101	0.708615
0.897748	1.141644	1.044462	1.173262	1.149322	0.958402	0.985481	0.797678
0.934498	0.74333	0.983117	1.018874	1.082834	0.958141	0.831689	1.080523
0.808465	0.840022	0.951164	1.006018	1.025462	1.023455	0.833135	1.128314

DZ	EA	EB	EC	ED	EE	EF	EG	EH	EI	EJ	EK
치과_trend	치약_trend	치킨_trend	카레_trend	칼국수_trend	커피_trend	컴퓨터_trend	콩나물_trend	탕수육_trend	택배_trend	통신요금_trend	파스타_trend
1.055429	0.978332	1.084506	0.910523	1.290773	0.969179	1.118842	1.174998	0.876697	1.040052	1.168148	0.988226
0.9641	0.906557	1.151388	0.713976	0.993488	1.0283	1.180062	0.974637	0.987117	1.102918	1.107275	1.013112
1.02565	0.983228	1.184413	0.933988	0.95592	0.992724	1.154351	1.037304	1.114922	1.106396	1.076262	1.022963
0.945217	1.000308	1.32083	0.890158	0.932581	1.000479	1.021791	0.931766	1.006157	1.069344	0.960808	1.113969

부록5 점유율 반영



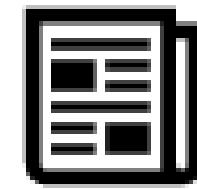
점유율

(1.25 x 0.16



점유율

+ 1.18 x 0.31



점유율

+ 1.32 x 0.53)

A	B	C	D	E	F	G	H	I	J
Unnamed: 0	ott_trend	pc방_trend	tv_trend	가구_trend	가방_trend	간장_trend	갈치_trend	감자_trend	계란_trend
2024-01-31 00:00:00	0.991541	1.033214	1.069663	0.983877	0.924923	0.987578	0.994528	0.928668	0.933465
2024-02-29 00:00:00	0.97811	1.038591	1.047935	1.07886	1.030158	1.346868	1.435601	1.11559	1.006451
2024-03-31 00:00:00	0.968315	0.896512	0.647507	0.991819	1.087175	1.024817	1.346877	0.987823	1.239674
2024-04-30 00:00:00	0.47749	0.916833	0.523712	1.021801	1.051992	0.953588	1.557965	0.828443	1.381457

부록6 개인별 소비 바스켓

	A	B
1	보험	8.6
2	지갑	0.5
3	우산	0.2
4	가방	2.7
5	시계	0.6
6	휴지	1.3
7	샴푸	1
8	비누	0.3
9	치약	0.9
10	이발	0.3
11	기숙사	0.9
12	도시락	0.6
13	소주	4.2
14	맥주	7.8

	A	B
1	부동산	0.8
2	보험	12.3
3	헤어샵	9.5
4	화장품	8.9
5	장신구	1
6	샴푸	1
7	바디워시	0.5
8	여행	11.9
9	커피	11.4
10	파스타	1.5
11	김밥	3.4
12	떡볶이	3.1
13	학원	27.6
14	학습지	3.5

	A	B
1	품목	가중치
2	요양	0.9
3	비누	0.3
4	찜질방	0.6
5	보험	8.6
6	샴푸	1
7	치약	0.9
8	비누	0.3
9	염색약	0.7
10	도시락	0.6
11	막걸리	0.3
12	맥주	4.8
13	소주	2.5
14	커피	2.6