

경희대학교 생활협동조합 카페 메뉴 주천시스템 개발

주천시스템 2차 프로젝트

주천시스템: 권정혁 도유정 최보경 최용빈

목차

1 프로젝트 개요

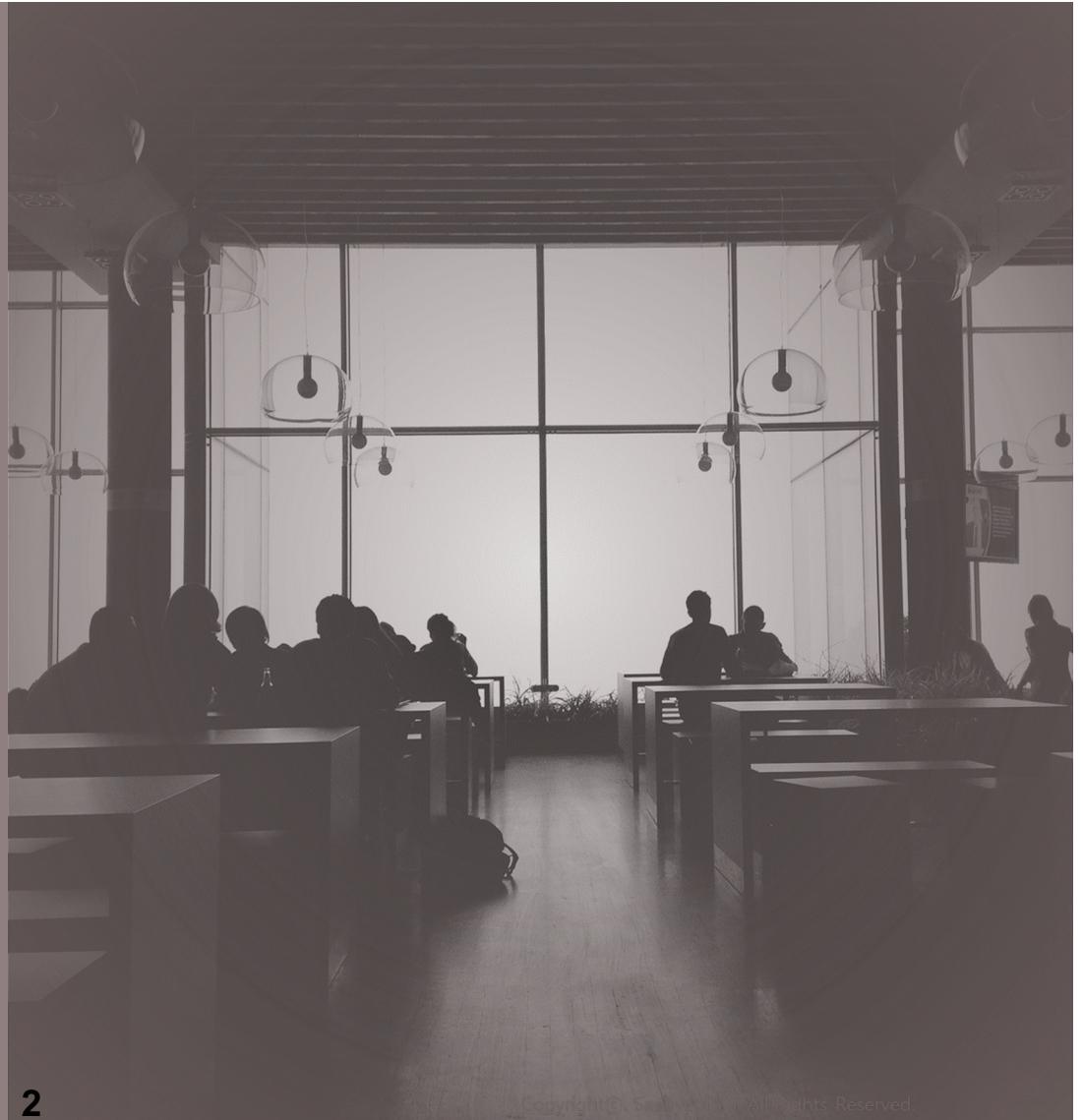
- 경희대학교 생협 소개

2 데이터 현황 및 문제점

3 모델링

- 유사도 측정
- 웹 만들기

4 결론 및 한계점



1

프로젝트 개요



1. 경희대학교 생협 카페란?



대학커피 ?

- 경영대 3층에 위치한 생협(생활협동조합) 카페
- 2020년 브랜드 '대학커피'로 새롭게 런칭
- 학생들을 위해 최대한 합리적인 가격으로 구성
- But 교내 제한된 공간으로 운영되어, 재고 관리가 어려움



1. 프로젝트 개요

프로젝트 목적	본 프로젝트는 카페데이터를 통해 재고관리의 문제점을 찾고, 이에 대한 해결방안을 제시한다.
프로젝트 기간	2023.05.20~2023.05.30





1. 프로젝트 개요

프로젝트 목적	본 프로젝트는 카페데이터를 통해 재고관리의 문제점을 찾고, 이에 대한 해결방안을 제시한다.
프로젝트 기간	2023.05.20~2023.05.30

카페 데이터 분석

문제점 분석

데이터 추가

추천시스템

- 데이터 전처리
- 재고관리 문제점 파악

- 메뉴관리 어려움
- 메뉴 조절 필요

- 날씨 데이터
- 음료특성 데이터

- 유사도분석



1. 프로젝트 개요

프로젝트 목적 본 프로젝트는 카페데이터를 통해 재고관리의 문제점을 찾고, 이에 대한 해결방안을 제시한다.

프로젝트 기간 2023.05.20 ~ 2023.05.30

날씨별 음료를 추천한다.

=> 이를 통해, 고객들이 기존 다양한 음료선택보다,
추천된 음료를 소비할 것으로 기대한다.

- 데이터 전처리
- 재고관리 문제점 파악

- 메뉴관리 어려움
- 메뉴 조절 필요

- 날씨 데이터
- 음료특성 데이터

- 유사도분석



1. 프로젝트 개요

프로젝트 목적 본 프로젝트는 카페데이터를 통해 재고관리의 문제점을 찾고, 이에 대한 해결방안을 제시한다.

프로젝트 기간 2023.05.20~2023.05.30

=> 추천시스템을 통해, 선택의 폭을 줄이는 것에 초점

카페 데이터 분석

문제점 분석

데이터 추가

추천시스템

- 데이터 전처리
- 재고관리 문제점 파악

- 메뉴관리 어려움
- 메뉴 조절 필요

- 날씨 데이터
- 음료특성 데이터

- 유사도분석



1. 프로젝트 개요



메뉴 관리의 필요성

날씨데이터

특성데이터



매출순위 고려

items	has_coffee	has_milk
아메리카노	1	0
아이스 카페라떼	1	1
아이스 바닐라라떼	1	1
카페라떼	1	1
아이스티북숭아	0	0

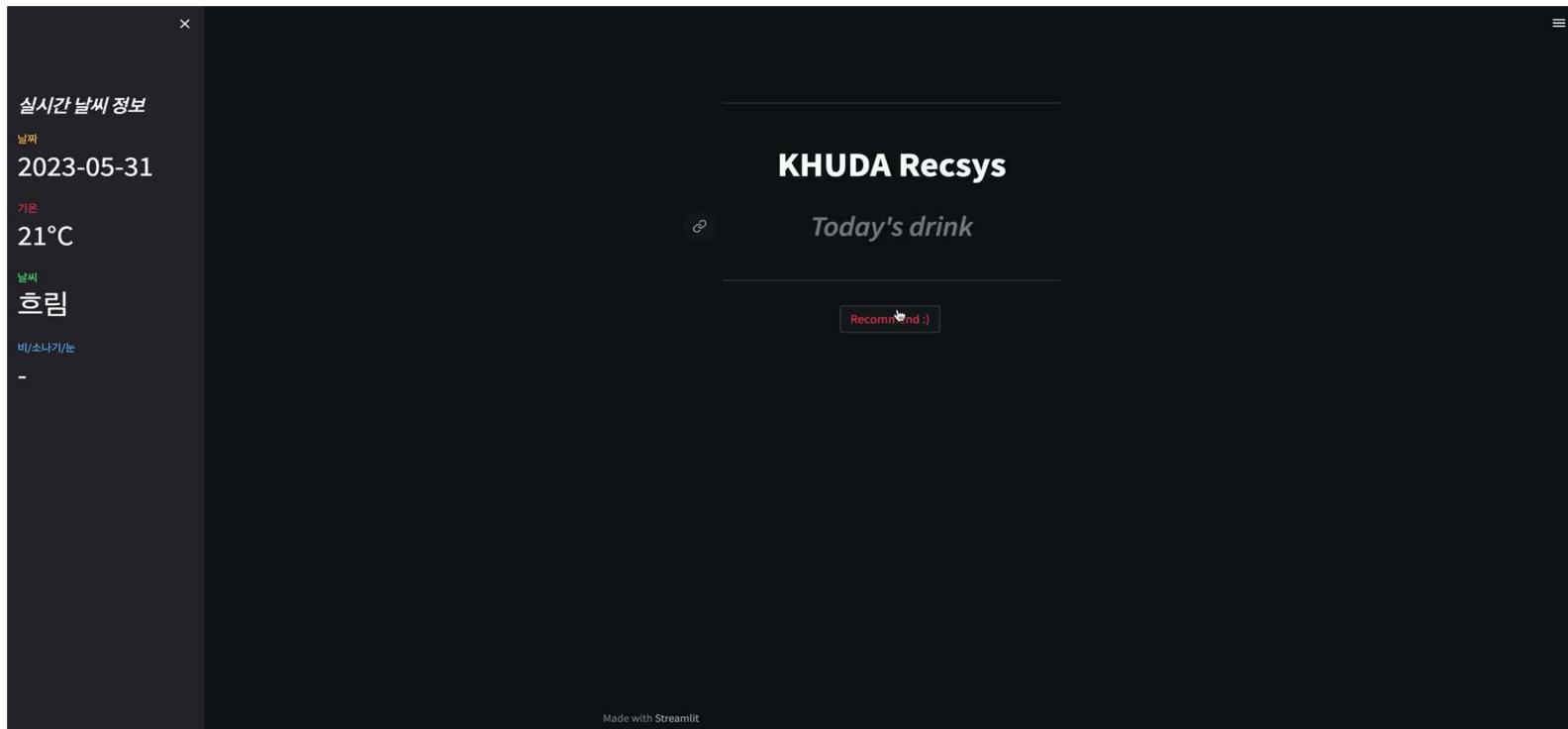


‘날씨별 메뉴 추천’ 제안



1. 프로젝트 개요

[웹 실행 영상]



2 데이터 현황 및 문제점



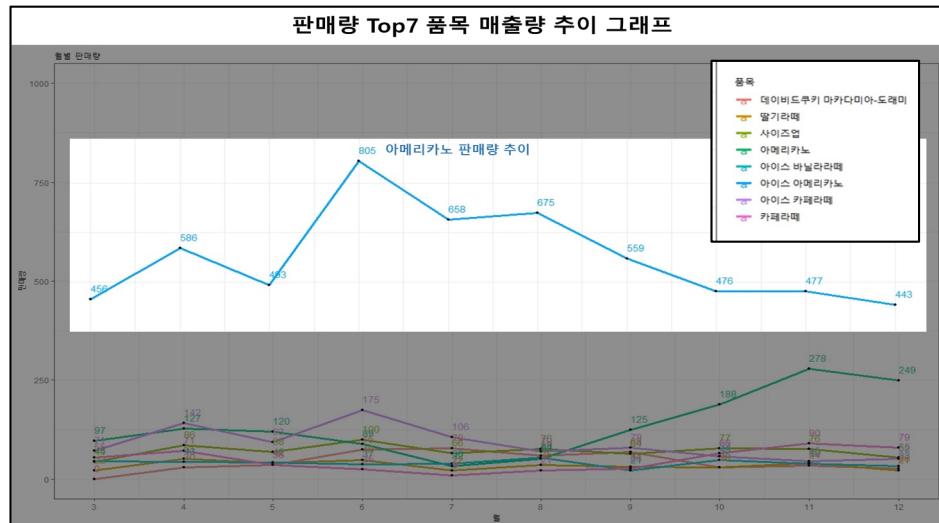
2. 데이터 현황

매출일시	품번	바코드	품명	수량	판매단가	총매출액	할인	실매출액	전표번호	영수증금액	할인명	메모	고객코드	고객명	형태	원매출일
2021-03-02 09:18:04	0002820	23000193	아이스 아메리카노	1	2,500	2,500	0	2,500	001-20210302-1	2,500					정상	
2021-03-02 09:43:41	0002820	23000193	아이스 아메리카노	1	2,500	2,500	0	2,500	001-20210302-2	6,000					정상	
2021-03-02 09:43:41	0002828	23000278	아이스 바닐라라떼	1	3,500	3,500	0	3,500	001-20210302-2						정상	
2021-03-02 09:44:03	0002820	23000193	아이스 아메리카노	-1	2,500	-2,500	0	-2,500	001-20210302-3	-6,000					반품	20210302
2021-03-02 09:44:03	0002828	23000278	아이스 바닐라라떼	-1	3,500	-3,500	0	-3,500	001-20210302-3						반품	20210302
2021-03-02 09:45:58	0002819	23000186	아메리카노	1	2,500	2,500	0	2,500	001-20210302-4	5,500					정상	
2021-03-02 09:45:58	0002823	23000223	카페라떼	1	3,000	3,000	0	3,000	001-20210302-4						정상	
2021-03-02 09:46:13	0002825	23000247	캬푸치노	1	3,000	3,000	0	3,000	001-20210302-5	3,000					정상	

- 2021-03-02 ~ 2022-08-25까지 1년 6개월간의 매출 데이터
- 약 26000개의 판매 데이터 수집



2. 데이터 현황



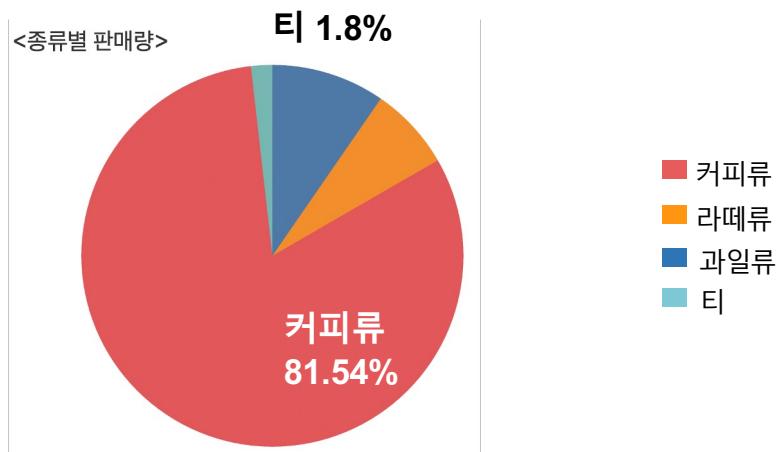
1년간 매출량 추이를 분석했을 때 계절에 상관없이
압도적으로 높은 판매량을 자랑하는 아메리카노



전체 판매량 중 압도적으로 높은
아이스 아메리카노



2. 데이터 현황



- 왼쪽 차트에서 볼 수 있듯이 전체적으로 편향된 판매
량을 확인 할 수 있음



2. 문제점 -1

```
print(len(df3['품명'].unique()))
print(df3['품명'].unique())
```

```
85
['아이스 아메리카노', '아이스 바닐라라떼', '아메리카노', '카페라떼', '카푸치노',
'모과차', '아이스티복숭아', '콜드브루라떼블랜딩', '콜드브루블랜딩', '사이즈',
'아이스 그린티라떼', '자몽에이드', '아이스 카라멜마끼아또', '아이스 밀크티',
'유자차', '초코라떼', '아이스 오폭라떼', '크리스틴 캐모마일', '민트초코라떼',
'크리스틴 키르로얄', '아이스 카페모카', '시럽추가', '콜드브루-도래미', '오곡',
'크리스틴 루이보스', '크리스틴 아이스 루이보스', '자몽차', '딸기라떼', '아이스',
'청포도에이드', '에스프레소', '휘핑추가', '크리스틴 아이스캐모마일', '스팀두',
'허니버터맛아몬드-도래미', '데이비드쿠키 마카다미아-도래미', '모카번-도래미',
'카라멜마끼아또', '크리스틴 얼그레이', '군옥수수맛아몬드-도래미', '사과에이드',
'초코머핀-도래미', '콜드브루싱글', '콜드브루 싱글오리진-판매', '크리스틴',
'콜드브루라떼싱글', '피지워터', '크리스틴 아이스루이보스', '딸기에이드', '아이스',
'오렌지에이드', '슈가와플-도래미', '치즈스콘-도래미', '크로크무슈-도래미',
'콤부차석류-도래미', '크리스틴 아이스키르로얄', '아인슈페너 콜드브루라떼',
'텐더쿠키모카-도래미']
```

- 메뉴가 너무 많음.
- 18개월동안 약 85개 메뉴 존재

=> 재고관리의 어려움



2. 문제점 - 2

	품명	판매량
36	아이스 카푸치노	30
67	크리스틴 루이보스	30
65	콤부차오리지널-도레미	33
56	카페모카	37
44	오렌지에이드	38
53	치즈스콘-도레미	39
76	크리스틴 얼그레이	40
40	에그타르트-도레미	41
84	휘핑추가	41
5	딸기에이드	41
70	크리스틴 아이스 얼그레이	42

- 판매기간동안 각 제품의 총판매 갯수
- 100개 미만인 제품 약 55개 => 재고관리의 어려움
- ‘잘팔리는 메뉴’에 집중할 필요

```
filtered_df_sales = sorted_df_sales[sorted_df_sales['판매량'] < 100]
len(filtered_df_sales)
```

55

3 모델링



3. 모델링

items	is_drink	has_coffee	has_milk	caffeine	is_carbonated	is_hot	has_choco	is_syrup_based	is_tea_variant	sugar	sugar(integer)	has_fruit	is_powdered	ig_hashtag	price	UI_click	time	calories	Tot.sales
그린티라떼	1	0	1	0.64615	0	1	0	0	0	1	0.63636	0	1	0.0316	0.05556	0.8	0.63636	0.64353	0.021892
딸기라떼	1	0	1	0	0	0	0	1	0	1	0.8409	1	0	0.11103	0.05556	0.4	0.22727	0.36585	0.237295
딸기에이드	1	0	0	0	1	0	0	1	0	1	0.68182	1	0	0.01998	0.08333	0.4	0.27273	0.21764	0.016028
레몬에이드	1	0	0	0	1	0	0	1	0	1	0.65909	1	0	0.08439	0.08333	0.4	0.27273	0.2758	0.154418
레몬차	1	0	0	0	0	1	0	1	1	1	0	1	0	0.01586	0.05556	0.8	0.22727	0.35835	0.02932
모과차	1	0	0	0	0	1	0	1	1	1	0.95456	1	0	0.00593	0.05556	0	0.22727	0.35647	0.021501
민트초코라떼	1	0	1	0.16154	0	1	1	1	0	1	0.72727	0	1	0.00448	0.05556	0	0.63636	1	0.004691
밀크티	1	0	1	0.26923	0	1	0	0	0	1	0.61364	0	1	0.27836	0.05556	0.8	0.63636	0.46717	0.032447
바닐라라떼	1	1	1	0.79231	0	1	0	1	0	1	0.40909	0	1	0.20668	0.08333	0.6	0.72727	0.53096	0.063331

—> 빨간색으로 표시된 행은 음료 특성 데이터입니다.

is_syrup_based	시럽 함유 여부
is_tea_variant	차 여부
sugar	당 함량 (integer)
has_fruit	과일 함유 여부
is_powdered	분말형 재료 사용 여부
ig_hashtag	인스타그램 해시태그 개수 (integer)
price	가격 (integer)
UI_click	UI 클릭 수(integer)
time	제조 시간 (integer)
calories	칼로리 (integer)
Tot.sales	총 판매량 (integer)

음료 특성 데이터

- 각 음료별 특성 데이터 수집
- 재료가 함유되어있으면 1, 아니면 0으로 나타낼 수 있는 feature들만 남기고 나머지는 제외
- 아이스 아메리카노 제외
=> 추천 결과가 편중될 가능성이 있음



3. 모델링

```
tmp['평균기온(°C)'].describe()  
count      541.00000  
mean       15.515157  
std        10.147891  
min       -12.100000  
25%        8.400000  
50%       17.900000  
75%       23.900000  
max       31.700000  
Name: 평균기온(°C), dtype: float64
```

- 기온(0~3), 운량(0~1), 강수여부(0~1) 기준으로 날씨를 16개의 범주로 그룹화
- 매출 데이터가 존재하는 날짜의 기온, 운량, 강수 여부 데이터 수집 후 그룹별로 분류
- 해당되는 데이터가 과하게 적은 그룹은 제거하려 했으나 존재하지 않아 그대로 진행



3. 모델링

기온	흐림	비소나기날
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0
1	1	1
2	0	0
2	0	1
2	1	0
2	1	1
3	0	0
3	0	1
3	1	0
3	1	1

총합
2235
967
516
729
3799
559
694
1486
2541
790
346
3105
1673
700
1249
2873

- 기온, 운량, 강수여부 기준으로 날씨를 12개의 범주로 그룹화
- 매출 데이터가 존재하는 날짜의 기온, 운량, 강수여부 데이터 수집 후 그룹별로 분류
- 해당되는 데이터가 과하게 적은 그룹은 제거하려 했으나 존재하지 않아 그대로 진행



3. 모델링

기온	흐림	비/소나기/눈	아이스 아메리카노	아메리카노	아이스 카페라떼	아이스 바닐라라떼	카페라떼	딸기라떼	레몬에이드	복숭아아이스티
0	0	0	0.4689038031	0.1821029083	0.051901566	0.03713646532	0.05369127517	0.02147651007	0.009843400447	0.004474272931
0	0	1	0.4291623578	0.1923474664	0.04033092037	0.04343329886	0.07962771458	0.02998965874	0.02481902792	0
0	1	0	0.4728682171	0.1511627907	0.04651162791	0.04651162791	0.05426356589	0.02131782946	0.02131782946	0.02519379845
0	1	1	0.4417009602	0.2249657064	0.04252400549	0.02606310014	0.0548696845	0.01920438957	0.01646090535	0
1	0	0	0.5022374309	0.1245064491	0.06528033693	0.04580152672	0.04843379837	0.02474335351	0.01263490392	0.01947881021
1	0	1	0.4186046512	0.1484794275	0.05903398927	0.04651162791	0.05903398927	0.02325581395	0.03220035778	0.005366726297
1	1	0	0.5187319885	0.1412103746	0.05187319885	0.03458213256	0.04034582133	0.01585014409	0.0144092219	0.02449567723
1	1	1	0.4703903096	0.1379542396	0.07537012113	0.03903095559	0.04172274563	0.03028263795	0.0067294751010	0.009421265141

items	has_coffee	has_milk	caffeine	is_carbonated	is_hot	has_choco	is_syrup	b_is_tea	variety	sugar	has_fruit	is_powdered
아메리카노	1	0	1	0	1	0	0	0	0	0	0	0

각 메뉴 판매량
날씨 그룹별 메뉴 판매량 총합

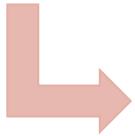
각 메뉴의 가중치

가중치를 메뉴별 특성에 곱해주기



3. 모델링

날씨 그룹별로 가중치가 곱해진 특성들을 특성별로 다 더해줌



기온	흐림	비소나기는	has_coffee	has_milk	caffeine	is_carbonated	is_hot	has_chocolate	is_syrup_b	is_tea_var
0	0	0	0.73902	0.483953	0.822635	0.046453	0.581926	0.058277	0.289696	0.054054
0	0	1	0.733696	0.528986	0.831522	0.048913	0.617754	0.04529	0.297101	0.043478
0	1	0	0.69403	0.511194	0.776119	0.067164	0.496269	0.052239	0.309701	0.033582
0	1	1	0.753695	0.445813	0.832512	0.051724	0.633005	0.054187	0.270936	0.046798
1	0	0	0.664896	0.520446	0.734997	0.077005	0.450876	0.035582	0.329262	0.052045
1	0	1	0.633846	0.461538	0.686154	0.153846	0.44	0.024615	0.369231	0.073846
1	1	0	0.636364	0.466667	0.721212	0.081818	0.509091	0.048485	0.312121	0.063636
1	1	1	0.657727	0.535121	0.752235	0.061303	0.457216	0.054917	0.320562	0.057471
2	0	0	0.605858	0.531381	0.676987	0.103766	0.257741	0.03431	0.364854	0.041841
2	0	1	0.573407	0.581717	0.648199	0.088643	0.221607	0.024931	0.454294	0.044321
2	1	0	0.666667	0.616352	0.72327	0.125786	0.301887	0.037736	0.36478	0.031447
2	1	1	0.62901	0.543236	0.711994	0.098326	0.316597	0.043236	0.352162	0.055091
3	0	0	0.568871	0.567493	0.658402	0.173554	0.152893	0.044077	0.436639	0.03719
3	0	1	0.560811	0.594595	0.64527	0.152027	0.148649	0.057432	0.483108	0.043919
3	1	0	0.564402	0.45666	0.615222	0.186047	0.238901	0.023256	0.376321	0.059197
		1	0.594178	0.560788	0.674658	0.117295	0.22089	0.037671	0.365582	0.057363

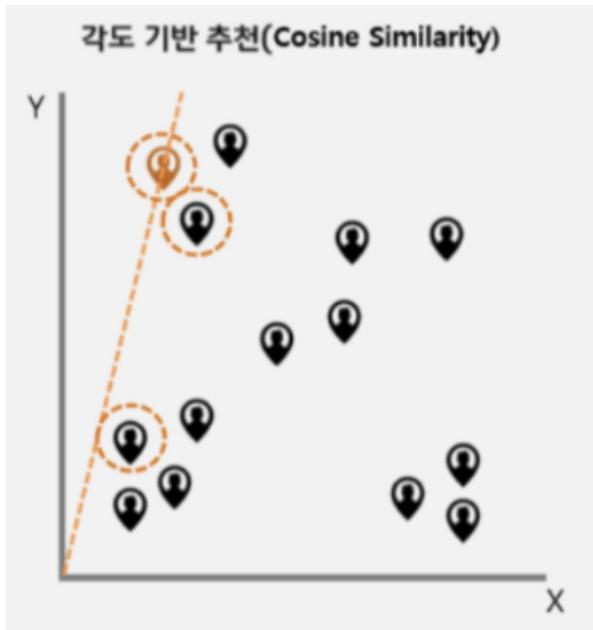
인기있던 특성일수록 1

에 수렴





3. 유사도 측정 (1) Cosine Similarity



- 코사인 유사도는 두 벡터가 가르키는 방향이 얼마나 유사한지에 따라 유사도를 측정
- 코사인 유사도는 -1에서 1 사이의 값을 가지며 1에 가까울수록 유사도가 높다고 할 수 있음



3. 유사도 측정 (1) Cosine Similarity

날씨
그룹

	아메리 카노	아이스 카페라 떼	아이스 바닐라 라떼	카페라 떼	아이스 티복숭 아	딸기라 떼	레몬에 이드	자몽에 이드	콜드브 루블랜 딩	청포도 에이드
a1	0.810193	0.865310	0.851439	0.944326	0.379878	0.499558	0.356353	0.356353	0.722902	0.356353
a2	0.799740	0.865600	0.852711	0.949520	0.387592	0.513268	0.360954	0.360954	0.702297	0.360954
a3	0.755252	0.878710	0.883655	0.933583	0.442833	0.556020	0.408327	0.408327	0.691550	0.408327
a4	0.840651	0.847612	0.828640	0.943865	0.352682	0.462187	0.332904	0.332904	0.735906	0.332904
a5	0.720813	0.883970	0.889949	0.926667	0.471993	0.600419	0.450851	0.450851	0.667745	0.450851
a6	0.709950	0.866358	0.868864	0.912376	0.472886	0.630958	0.523470	0.523470	0.652131	0.523470
a7	0.749831	0.861263	0.872782	0.928741	0.463776	0.573472	0.439592	0.439592	0.667892	0.439592
a8	0.724041	0.886088	0.890177	0.929874	0.464040	0.592727	0.433609	0.433609	0.669624	0.433609
a9	0.619378	0.893899	0.918041	0.879793	0.539980	0.676251	0.527365	0.527365	0.631670	0.527365
a10	0.557743	0.873330	0.933350	0.847468	0.572731	0.724993	0.559969	0.559969	0.578204	0.559969

메뉴

그룹 기준

COS 유사도

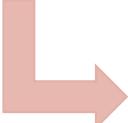
계산 결과



3. 유사도 측정 (1) Cosine Similarity

```
top_5_cols = pd.DataFrame()  
  
for i in range(1,17):  
    top_5 = k1.loc['a'+str(i)].nlargest(5).index  
    top_5_cols['a'+str(i)] = top_5
```

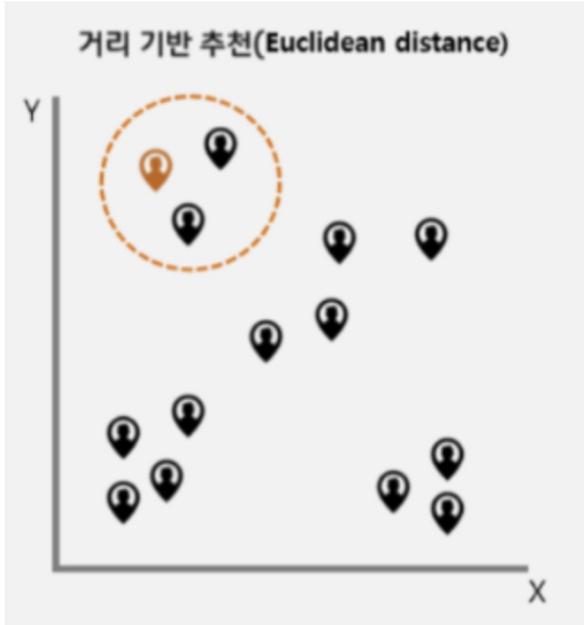
그룹별 코사인
유사도가 높은
메뉴 5개 추출



	0	1	2	3	4
a1	카페라떼	바닐라라떼	카푸치노	아이스 카페라떼	콜드브루라떼블랜딩
a2	카페라떼	바닐라라떼	카푸치노	아이스 카페라떼	콜드브루라떼블랜딩
a3	바닐라라떼	카푸치노	카페라떼	아이스 바닐라라떼	아이스 카페라떼
a4	카페라떼	카푸치노	바닐라라떼	아이스 카페라떼	콜드브루라떼블랜딩
a5	바닐라라떼	카페라떼	카푸치노	아이스 카라멜마끼아또	아이스 바닐라라떼
a6	바닐라라떼	카페라떼	아이스 카라멜마끼아또	카푸치노	아이스 바닐라라떼
a7	바닐라라떼	카페라떼	카푸치노	아이스 바닐라라떼	아이스 카라멜마끼아또
a8	바닐라라떼	카페라떼	카푸치노	아이스 바닐라라떼	아이스 카라멜마끼아또
a9	아이스 바닐라라떼	바닐라라떼	아이스 카라멜마끼아또	아이스 카페라떼	콜드브루라떼블랜딩
a10	아이스 바닐라라떼	바닐라라떼	아이스 카라멜마끼아또	아이스 카페라떼	콜드브루라떼블랜딩



3. 유사도 측정 (2) Euclidean distance



- 거리 기반 유사도는 좌표를 기준으로 봤을 때, 비슷한 좌표에 있는 점들이 유사도가 높다고 측정
- 유클리디안 거리는 N차원 공간에서 두 점 사이의 최단 거리를 구하기 위한 가장 기초적인 알고리즘임



3. 유사도 측정 (2) Euclidean distance

최종 데이터의 형태

기 운 리 기 준	비 소 나 기 준	has_coffee	has_milk	caffeine	is_carbonated	is_hot	has_chocolate	
0	0	0	0.739020	0.483953	0.822635	0.046453	0.581926	0.058277

음료 특 성

↑ 유사도 측
↓ 정

items	has_coffee	has_milk	caffeine	is_carbonated	is_hot	has_chocolate
아메리카노	1	0	1	0	1	0
아이스 카페라떼	1	1	1	0	0	0
아이스 바닐라라떼	1	1	1	0	0	0
카페라떼	1	1	1	0	1	0
아이스티북숭아	0	0	0	0	0	0
딸기라떼	0	1	0	0	0	0
레몬에이드	0	0	0	1	0	0



3. 유사도 측정 (2) Euclidean distance

```
def euclidean_distance(x1, x2):
    return np.sqrt(np.sum((x1 - x2) ** 2))

def find_similarity(numpy_feature, drink_feature):
    answer = []
    for i in range(37):
        drink_feature_solo = drink_feature.iloc[i, 1:]
        numpy_drink_feature = drink_feature_solo.to_numpy()
        numpy_drink_feature = numpy_drink_feature.astype(float)
        value = euclidean_distance(numpy_feature, numpy_drink_feature)
        answer.append(value)
    return answer

def get_min_indices(lst, n):
    min_indices = []
    sorted_indices = sorted(range(len(lst)), key=lambda i: lst[i])
    for i in range(n):
        min_indices.append(sorted_indices[i])
    return min_indices

def get_items(min_indices):
    min_items = []
    for i in range(len(min_indices)):
        min_items.append(drink_feature.loc[min_indices[i], 'items'])
    return min_items

def recommend_drink_by_weather(feature_weather, drink_feature):
    lst = []
    for i in range(len(feature_weather)):
        numpy_feature = feature_weather.iloc[i]
        value = find_similarity(numpy_feature, drink_feature)
        lst.append(value)
    return lst
```

- 유클리디안 거리를 구한 뒤 거리가 제일 가까운 n개의 음료를 추천해주는 함수구현



3. 유사도 측정 (2) Euclidean distance

- 0 ['카페라떼', '아메리카노', '아이스 카페라떼', '콜드브루라떼블랜딩', '에스프레소']
- 1 ['카페라떼', '아이스 카페라떼', '콜드브루라떼블랜딩', '아메리카노', '에스프레소']
- 2 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '콜드브루블랜딩', '콜드브루싱글']
- 3 ['아메리카노', '에스프레소', '카페라떼', '콜드브루라떼블랜딩', '콜드브루싱글']
- 4 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아이스 카라멜마끼아또', '콜드브루블랜딩']
- 5 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아이스 카라멜마끼아또', '콜드브루블랜딩']
- 6 ['카페라떼', '아이스 카페라떼', '콜드브루라떼블랜딩', '아메리카노', '에스프레소']
- 7 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아이스 카라멜마끼아또', '콜드브루블랜딩']
- 8 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '카페라떼', '콜드브루블랜딩']
- 9 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '아이스 바닐라라떼', '아이스 밀크티']
- 10 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '카페라떼', '콜드브루블랜딩']
- 11 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '카페라떼', '콜드브루블랜딩']
- 12 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '아이스 바닐라라떼', '아이스 밀크티']
- 13 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '아이스 바닐라라떼', '아이스 밀크티']
- 14 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '콜드브루블랜딩', '콜드브루싱글']
- 15 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '카페라떼', '아이스 밀크티']

각 날씨별 유사도가 높은 5개의 메뉴
추출



날씨 그룹별 많이 팔린 음료 알 수
있음

But

평소보다 특정 날씨에 많이 팔린 음
료는 알 수 없음



#3 유사도 측정 (2) Euclidean distance

0 ['카페라떼', '아메리카노', '아이스 카페라떼', '콜드브루라떼블랜딩', '에스프레소']
1 ['카페라떼', '아이스 카페라떼', '콜드브루라떼블랜딩', '아메리카노', '에스프레소']
2 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아메리카노', '에스프레소']
3 ['아메리카노', '에스프레소', '콜드브루라떼블랜딩', '카페라떼', '아이스 카라멜마끼아또']
4 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아이스 카라멜마끼아또', '콜드브루라떼블랜딩']
5 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아이스 카라멜마끼아또', '콜드브루라떼블랜딩']
6 ['카페라떼', '아이스 카페라떼', '콜드브루라떼블랜딩', '아메리카노', '에스프레소']
7 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아메리카노', '에스프레소']
8 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아메리카노', '에스프레소']
9 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '아이스 바닐라라떼', '아이스 밀크티']
10 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '카페라떼', '콜드브루라떼블랜딩']
11 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '카페라떼', '콜드브루라떼블랜딩']
12 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '아이스 바닐라라떼', '아이스 밀크티']
13 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '아이스 바닐라라떼', '아이스 밀크티']
14 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '콜드브루라떼블랜딩', '콜드브루싱글']
15 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '카페라떼', '아이스 밀크티']

각 날씨별 유사도가 높은 5개의 메뉴
'상대적으로' 인기있던 특성을 기준
으로 추천해주면 어떨까?

각 날씨별 유사도가 높은 5개의 메뉴

날씨 그룹별 많이 팔린 음료 알 수
있음

But
평소보다 그 날씨에만 많이 팔린 음
료는 알 수 없음



3. 유사도 측정 (2) Euclidean distance

	has_coffee	has_milk	caffeine	is_carbonated	is_hot	has_chocolate
0	0.739020	0.483953	0.822635	0.046453	0.581926	0.058277
1	0.733696	0.528986	0.831522	0.048913	0.617754	0.045290
2	0.694030	0.511194	0.776119	0.067164	0.496269	0.052239



전체 평균을 기준으로 데이터 표준화
진행

	has_coffee	has_milk	caffeine	is_carbonated	is_hot	has_chocolate
0	0.594994	0.458353	0.600119	0.450810	0.704727	0.516174
1	0.589669	0.503386	0.609006	0.453271	0.740555	0.503186
2	0.550003	0.485594	0.553604	0.471522	0.619070	0.510135



3. 유사도 측정 (2) Euclidean distance

0 ['아메리카노', '에스프레소', '카페라떼', '카푸치노', '크리스틴 캐모마일']
1 ['아메리카노', '에스프레소', '카페라떼', '카푸치노', '바닐라라떼']
2 ['아메리카노', '에스프레소', '카푸치노', '카페라떼', '바닐라라떼']
3 ['아메리카노', '에스프레소', '크리스틴 캐모마일', '카페라떼', '콜드브루블랜딩']
4 ['아메리카노', '에스프레소', '카푸치노', '카페라떼', '크리스틴 캐모마일']
5 ['유자차', '레몬차', '모과차', '자몽차', '레몬에이드']
6 ['크리스틴 캐모마일', '아메리카노', '에스프레소', '유자차', '레몬차']
7 ['아메리카노', '에스프레소', '카푸치노', '카페라떼', '밀크티']
8 ['레몬에이드', '자몽에이드', '청포도에이드', '유자에이드', '딸기라떼']
9 ['딸기라떼', '아이스 레몬차', '아이스 유자차', '레몬에이드', '자몽에이드']
10 ['아이스 카라멜마끼아또', '아이스 카페라떼', '콜드브루라떼블랜딩', '딸기라떼', '레몬에이드']
11 ['아이스 오곡라떼', '딸기라떼', '아이스 초코라떼', '아이스 민트초코라떼', '아이스 밀크티']
12 ['레몬에이드', '자몽에이드', '청포도에이드', '유자에이드', '딸기라떼']
13 ['딸기라떼', '레몬에이드', '자몽에이드', '청포도에이드', '유자에이드']
14 ['레몬에이드', '자몽에이드', '청포도에이드', '유자에이드', '아이스 레몬차']
15 ['딸기라떼', '레몬에이드', '자몽에이드', '청포도에이드', '유자에이드']

각 날씨별 유사도가 높은 5개의 메뉴
추출



날씨 그룹별 평소보다 많이 팔린 음료 알
수 있음
But
평소에 많이 팔리는 음료가 고려되지 않음



3. 유사도 측정 (2) Euclidean distance

-
- 0 ['카페라떼', '아메리카노', '아이스 카페라떼', '에스프레소', '카푸치노', '크리스틴 캐모마일']
 - 1 ['카페라떼', '아이스 카페라떼', '콜드브루라떼블랜딩', '아메리카노', '에스프레소', '카푸치노']
 - 2 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아메리카노', '에스프레소', '카푸치노']
 - 3 ['아메리카노', '에스프레소', '카페라떼', '크리스틴 캐모마일', '콜드브루블랜딩', '콜드브루싱글']
 - 4 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아메리카노', '에스프레소', '카푸치노']
 - 5 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '유자차', '레몬차', '모과차']
 - 6 ['카페라떼', '아이스 카페라떼', '콜드브루라떼블랜딩', '크리스틴 캐모마일', '아메리카노', '에스프레소']
 - 7 ['아이스 카페라떼', '콜드브루라떼블랜딩', '카페라떼', '아메리카노', '에스프레소', '카푸치노']
 - 8 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '레몬에이드', '자동에이드', '청포도에이드']
 - 9 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '딸기라떼', '아이스 레몬차', '아이스 유자차']
 - 10 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '딸기라떼', '레몬에이드', '자동에이드']
 - 11 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '아이스 오곡라떼', '딸기라떼', '아이스 초코라떼']
 - 12 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '레몬에이드', '자동에이드', '청포도에이드']
 - 13 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '딸기라떼', '레몬에이드', '자동에이드']
 - 14 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '레몬에이드', '자동에이드', '청포도에이드']
 - 15 ['아이스 카페라떼', '콜드브루라떼블랜딩', '아이스 카라멜마끼아또', '딸기라떼', '레몬에이드', '자동에이드']
-

원래의 데이터에서 추천된 상위 3개의 음료와 표준화된 데이터에서
추천된 상위 3개의 음료를 합쳐 총 6개의 음료를 추천하기로 결정



3. 웹 만들기



파이썬에서 분석한 내용을 웹에 띄울 수 있음

```
# 실시간 날씨 정보 받아오기
village_open_api =
    'http://apis.data.go.kr/1360000/VillageFcstInfoService_2
    .0/getUltraSrtFcst?'
```

기상청 단기예보 API 활용
→ 실시간 날씨 받아오기

```
def recommend(temp, cloud, falling):
    by_weather = pd.read_csv('feature_by_weather_without_iceamericano.csv', encoding = 'cp949')
    weather_group = by_weather[(by_weather['기온'] == temp) & (by_weather['흐림'] == cloud) &
                                (by_weather['비소나기'] == falling)]
    weight_feature = weather_group.drop(['기온', '흐림', '비소나기'], axis=1).to_numpy()

    by_avg_weather = pd.read_csv('ratio_features_without_iceamericano.csv', encoding = 'cp949')
    avg_weather_group = by_avg_weather[(by_avg_weather['기온'] == temp) &
                                         (by_avg_weather['흐림'] == cloud) & (by_avg_weather['비소나기'] == falling)]
    avg_weight_feature = avg_weather_group.drop(['기온', '흐림', '비소나기'], axis=1).to_numpy()

    drink_feature = pd.read_csv('features_without_iceamericano.csv', encoding =
        'cp949').drop(['Unnamed: 0'], axis=1)

    result = [] # 유사도 계산 결과
    result_avg = [] # 유사도 계산 결과
```

날씨에 따른 추천 메뉴 받아오기



3. 웹 만들기

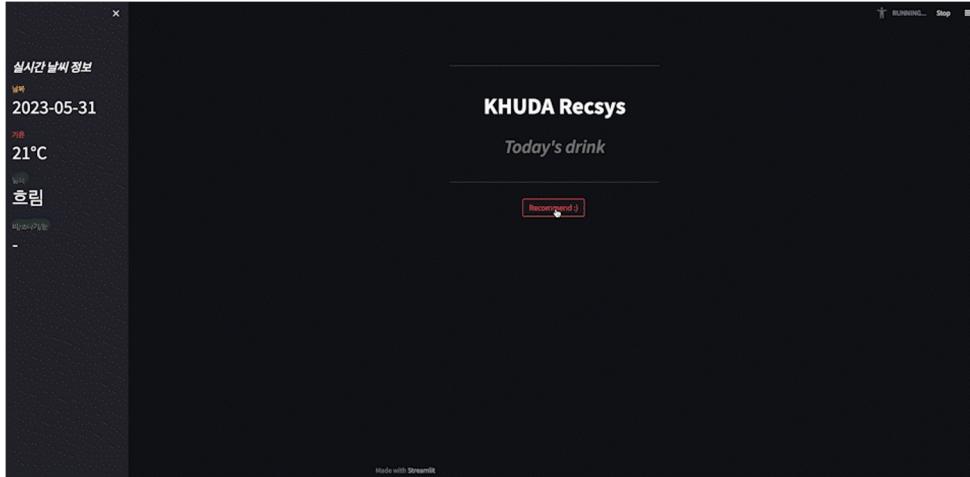
The screenshot shows a dark-themed web application interface. On the left, a sidebar displays real-time weather information: 날짜 (May 31, 2023), 기온 (21°C), 날씨 (흐림). A red box highlights this sidebar, and a red arrow points from it to a callout box containing the Korean text "실시간 날씨 정보를 보여주는 사이드바". The main content area is titled "KHUDA Recsys" and "Today's drink". Below this, there is a "Recommend" button. To the right, a large callout box contains the text "날씨에 따라 달라지 는 추천 메뉴 6개" (Recommendation menu changes according to the weather) in red. A red arrow points from the sidebar towards this callout. Further to the right, another callout box displays six recommended drink options, each with an image and a number: [1] 아이스 카페리아, [2] 풀드브루리페블렌딩, [3] 아이스 카라멜마끼아또, [4] 딸기리페, [5] 레몬에이드, [6] 자몽에이드.

4

결과 및 한계점



4. 결론

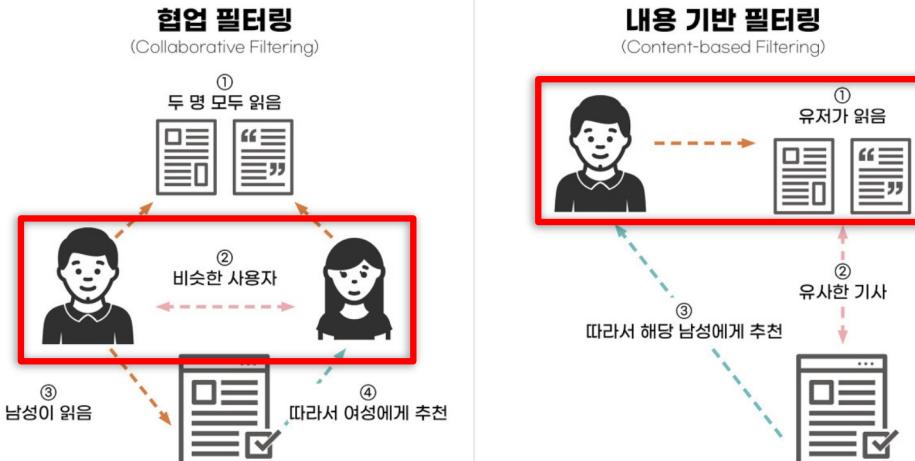


실시간 날씨를 받아와 메뉴를 추천
해 줌

- 추천 메뉴로 주문이 모이는 효과를
기대할 수 있음
→ 재고 관리 용이
- 고객의 입장에서 음료 선택의 고민
을 줄일 수 있음



4. 한계 (1) 유저 데이터 없음

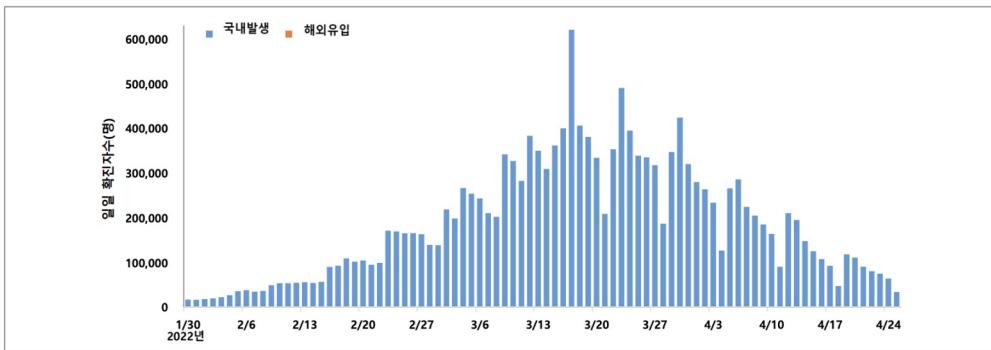


추천 시스템 : 사용자의 과거 행동 데이터나 다른 데이터를 바탕으로 사용자에게 필요한 정보나 제품을 골라서 제시해 주는 시스템

사용자에 대한 데이터가 없었음
→ 개인화 추천이 불가능함,
추천을 활용할 수 있는 선택지가 줄어들음



4. 한계 (2) 데이터의 부족



경영대카페의 매출데이터

2021.03.02 ~ 2022.08.25

- 코로나 시기라는 자료의 특이성
(코로나가 끝난 현재와의 결과가 달라질 수 있음)
- 유의미한 결과를 끌어내기에는 부족한 데이터의 양

Q&A