

# 효돌 **고객 군집화** 및 **관리 계열/시간대** 선정

( feat. **효린이** (고객관리 대시보드) )



Team : KLCP



# 목 차

---

1. 배경 소개 & 분석 목표

2. 데이터 소개

3. 고객 군집화 & 관리 계절/시간대 예측

## 3-1. 고객 군집화

- 1) 변수 생성 ( 요인 분석 )
- 2) 군집화 알고리즘 ( SOM )
- 3) **군집 별 페르소나**  
( + 행동 패턴/생활 양식 비교분석 )

## 3-2. 관리 계절/시간대 예측

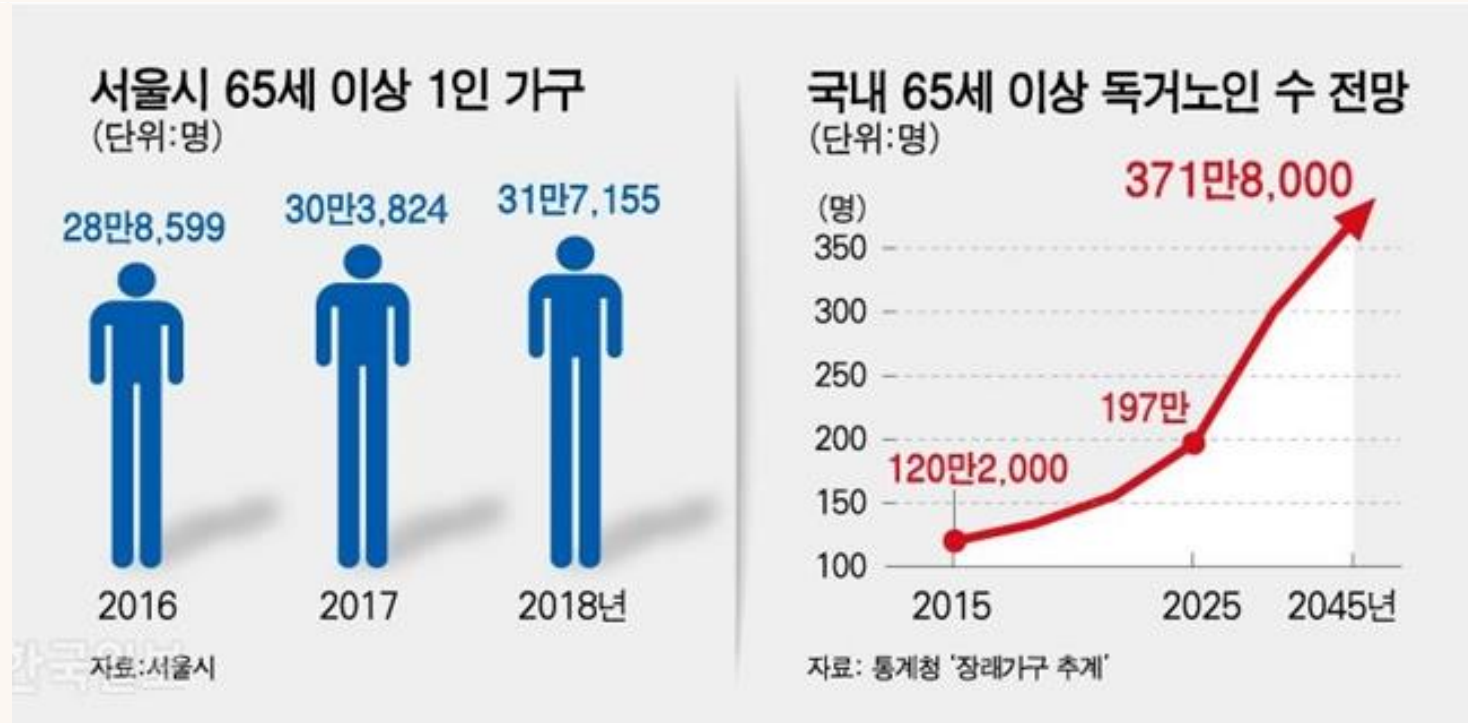
- 1) 응급 전화 내역 기반 예측
- 2) 월/시간대 예측 알고리즘 ( Multi-task Learning )
- 3) **'군집 별' 관리 계절/시간대 선정**

4. 군집 별 제언

# 1. 배경 소개 및 분석 목표

# 1. 배경 소개 및 분석 목표

## 1-1. 배경 소개



**증가하는 독거 노인**의 수...  
이러한 현실 속, **효돌의 역할**은?



[http://newsimg.hankookilbo.com/2018/10/22/201810221637063936\\_4.jpg](http://newsimg.hankookilbo.com/2018/10/22/201810221637063936_4.jpg)

# 1. 배경 소개 및 분석 목표

## 1-1. 배경 소개

<https://pds.joongang.co.kr/news/>

액티브 시니어와 전통노인의 생활의식 비교

구분	전통적 노년층	액티브 시니어
생활 의식	보수, 비관적인 인생관	합리적, 미래지향적
노년기의 의식	인생의 종말기	자기실현의 기회, 제3의 인생
삶의 태도	검약, 소박, 취미가 없음	여유, 즐길, 다양한 취미
독립성	자녀 등에 의지	강한 독립심, 사회 시스템에 의지
노후 설계	자녀세대에 의지	계획적인 노후설계
가치관	노인은 노인답게	나이와 젊음은 별개
레저관	일의 재미, 여가는 수단	여가 자체의 가치 목적화
자산 처분	자녀에 상속	본인을 위한 처분
여행 형태	친목 등 단체 여행	여유 있는 부부여행
취미 생활	노인끼리 교류	취미의 다양화, 다른 세대와 교류
생활 스타일	순 한국식 선호	타 문화와 교류하는 생활

[배경 1] 노년층의 **다양해진 욕구**

<https://lh3.googleusercontent.com/>

<응급안전알림서비스 응급상황 접수 현황>

(단위: 건)

	독거 노인			장애인		
	응급상황 신고건수	오작동 민감작동	출동건수	응급상황 신고건수	오작동 민감작동	출동건수
2015	163,665	60,651	2,199	8,549	2412	49
2016	147,780	53,265	1,913	7,010	2210	48
2017	135,661	55,964	1,358	10,728	3397	103
2018	110,609	44,615	1,697	13,873	2752	173

\*자료: 김상의 의원실(사외보장정보원 제출)

[배경 2] 응급 상황에서의 **미흡한 대처**

# 1. 배경 소개 및 분석 목표

## 1-1. 배경 소개

<https://pds.joongang.co.kr/news/>

액티브 시니어와 전통노인의 생활의식 비교

구분	전통적 노년층	액티브 시니어
생활 의식	보수, 비관적인 인생관	합리적, 미래지향적
노년기의 의식	인생의 종말기	자기실현의 기회, 제3의 인생
삶의 태도	검약, 소박, 취미가 없음	여유, 즐길, 다양한 취미
독립성	자녀 등에 의지	강한 독립심, 사회 시스템에 의지
노후 설계	자녀세대에 의지	계획적인 노후설계
가치관	노인은 노인답게	나이와 젊음은 별개
레저관	일의 재미, 여가는 수단	여가 자체의 가치 목적화
자산 처분	자녀에 상속	본인을 위한 처분
여행 형태	친목 등 단체 여행	여유 있는 부부여행
취미 생활	노인끼리 교류	취미의 다양화, 다른 세대와 교류
생활 스타일	순 한국식 선호	타 문화와 교류하는 생활

[배경 1] 노년층의 **다양해진 욕구**

➡ **[1] 노인 특징별 고객 군집화**

<https://lh3.googleusercontent.com/>

<응급안전알림서비스 응급상황 접수 현황>

(단위: 건)

	독거 노인			장애인		
	응급상황 신고건수	오작동 민감작동	출동건수	응급상황 신고건수	오작동 민감작동	출동건수
2015	163,665	60,651	2,199	8,549	2412	49
2016	147,780	53,265	1,913	7,010	2210	48
2017	135,661	55,964	1,358	10,728	3397	103
2018	110,609	44,615	1,697	13,873	2752	173

\*자료: 김상익 의원실(사외보장정보원 제출)

[배경 2] 응급 상황에서의 **미흡한 대처**

➡ **[2] 노인 별(군집 별) 관리 월/시간대 선정**

# 1. 배경 소개 및 분석 목표

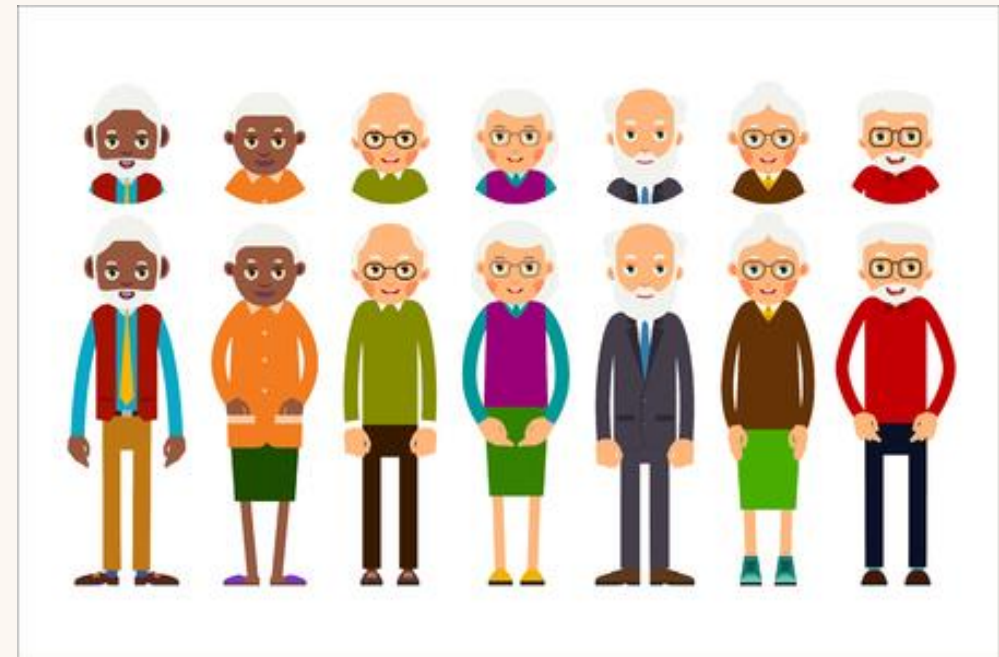
## 1-2. 분석 목표

노인의 생활 양식, 행동 패턴을 기반으로,

**1) 다양한 노인 군집**을 생성하고,

응급 상황 대처를 위한 **2) 노인 별 (+군집 별)**

**관리 계절/시간대**를 설정하자



<https://us.123rf.com/450wm/jeysent>

## 2. 데이터 소개



## 2. 데이터 소개

### 2-1. 사용한 데이터 소개

1

**인형/유저 종합 정보** : doll 정보, user 정보, agency 정보 table 합쳐서 생성

정적 정보

2

**귀 프로그램 수행 로그** : 이야기, 종교 말씀, 트로트, 영어 교실, 회상 등의 수행 횟수

3

**인형 접속 로그** : 인형 머리 쓰다듬, 손 버튼 누름, 등 두드림 ..

4

**응급 전화 내역 로그** : 인형 별 응급 전화 내역

동적(로그) 정보

## 2. 데이터 소개

### 2-2. 데이터 사용 흐름도

1

인형/유저 종합 정보

기본 정보

2

귀 프로그램 수행 로그

3

인형 접속 로그

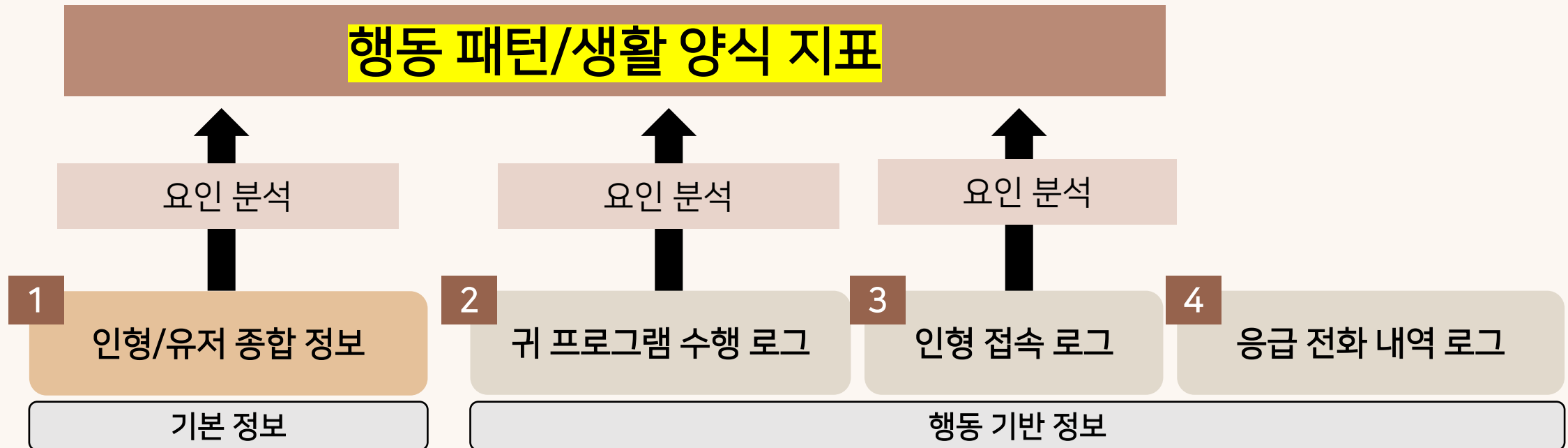
4

응급 전화 내역 로그

행동 기반 정보

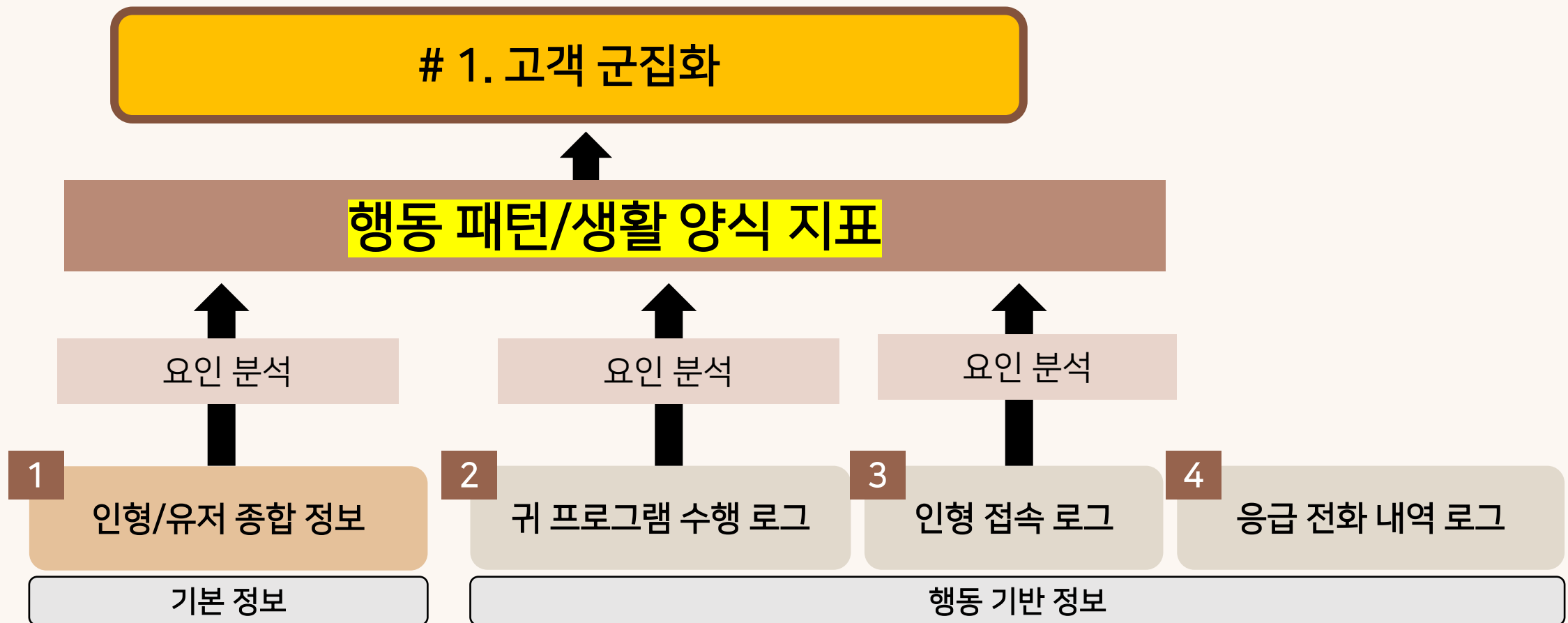
## 2. 데이터 소개

### 2-2. 데이터 사용 흐름도



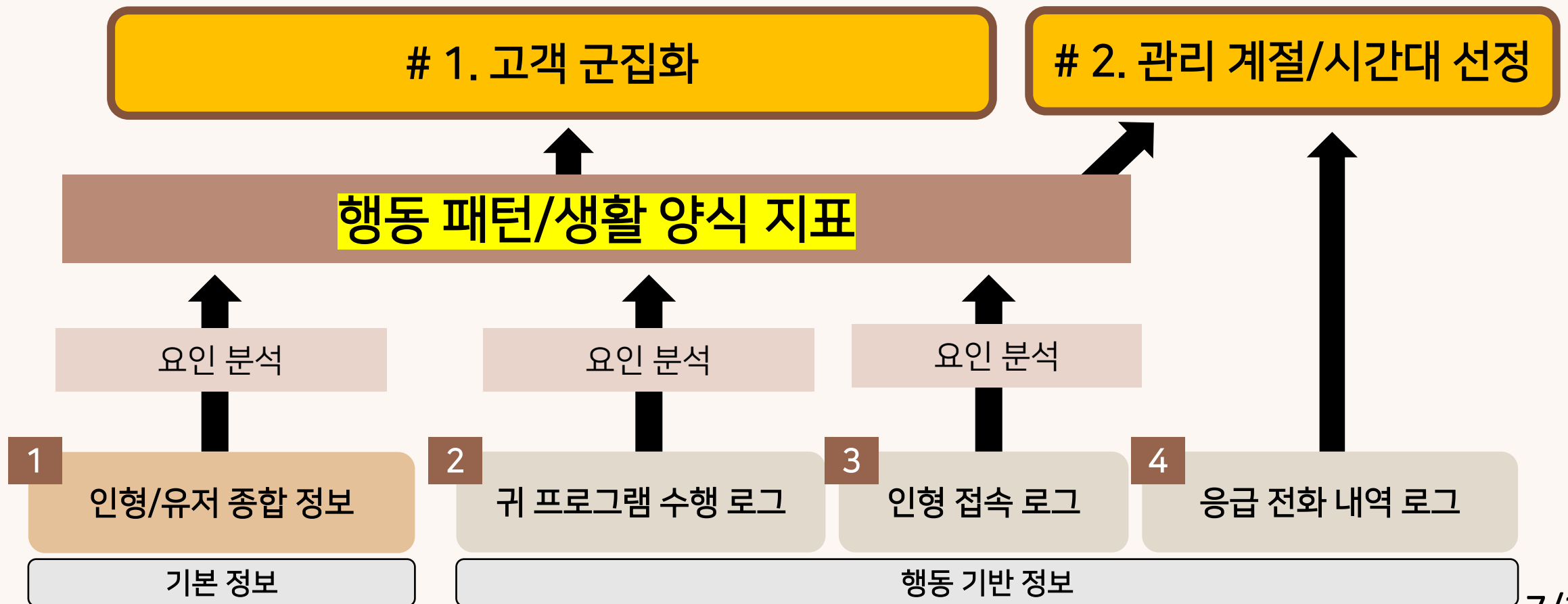
## 2. 데이터 소개

### 2-2. 데이터 사용 흐름도



## 2. 데이터 소개

### 2-2. 데이터 사용 흐름도



### 3. 고객 군집화 & 집중관리 시간대 예측

#### (1) 고객 군집화

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성

(인형 수 : 3126)

1

인형/유저 종합 정보

(인형 수 : 2504)

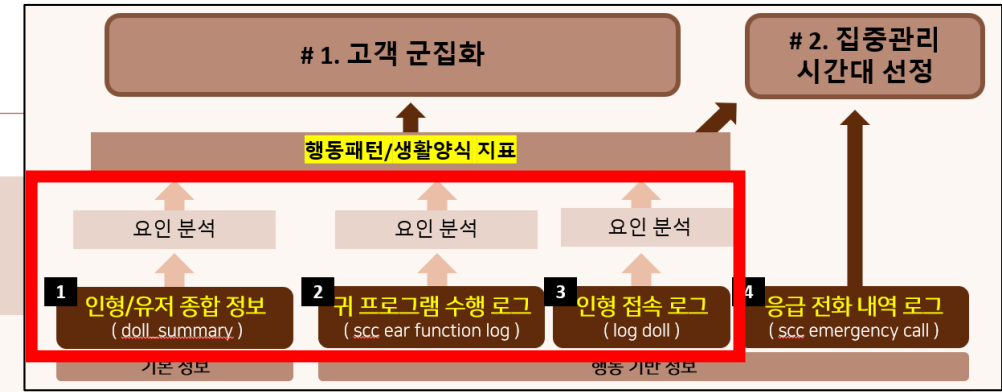
2

귀 프로그램 수행 로그

(인형 수 : 7981)

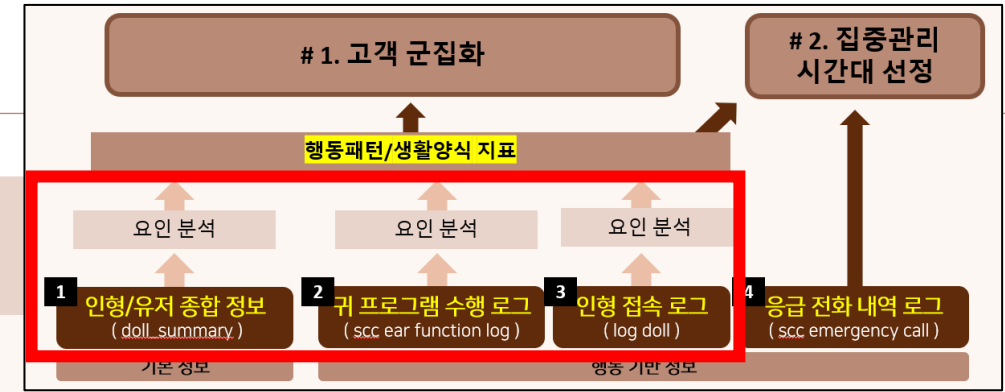
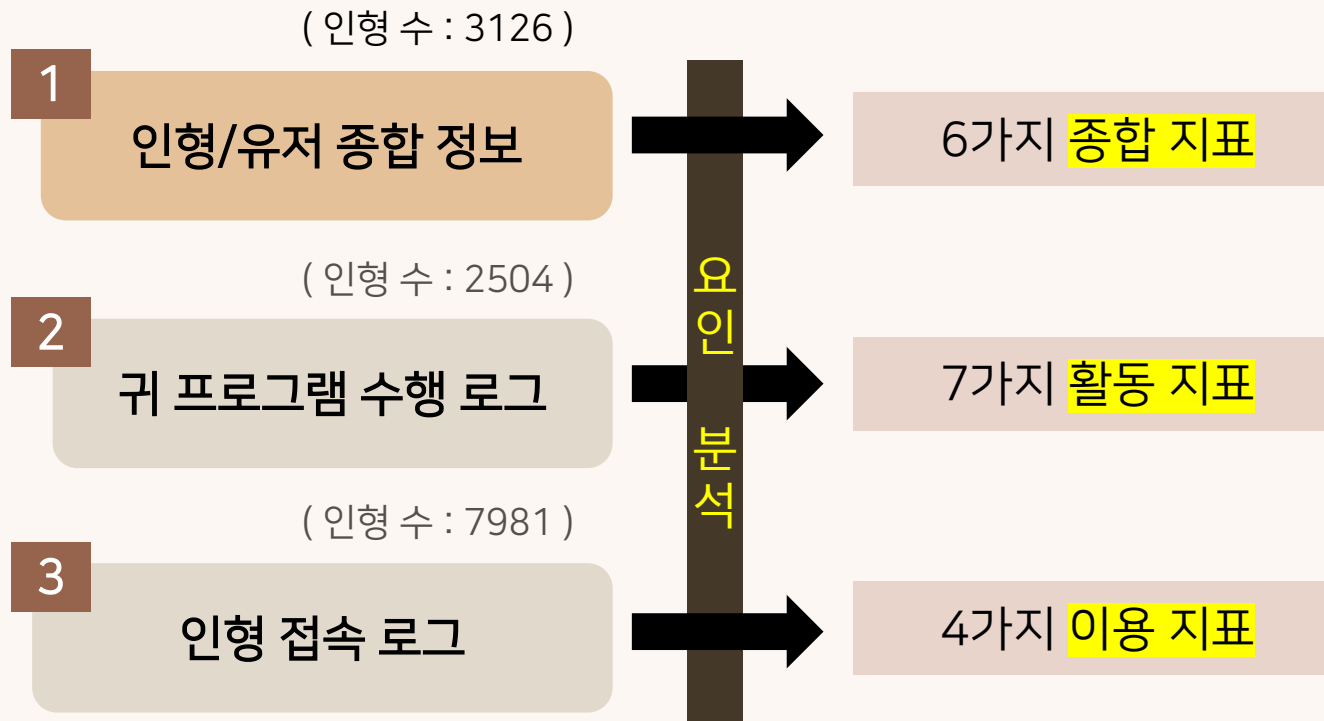
3

인형 접속 로그



## 3-1. 고객 군집화

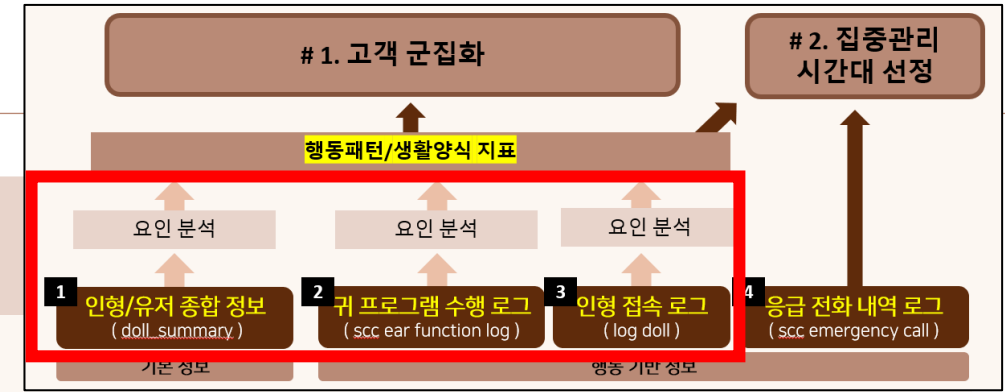
### (1) 군집화를 위한 변수 생성





## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성



\* 요인 분석 과정 공통 사항 :

1차 표준화를 하여하여, z-score를 계산한 뒤 이상치 선정 (-3~3 외의 범위), 이상치를 제외한 값들만으로 평균/표준편차 구한 뒤, 2차 표준화

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성

#### 17가지의 행동 패턴/생활 양식 지표

##### [ 6가지 종합 지표 ]

###### [# 1-1. 수면 패턴]

- 새벽 & 올빼미형

###### [# 1-2. 규칙적 생활]

- 규칙적 아침 & 점심/저녁형

###### [# 1-3. 약 복용]

- 식전 & 식후 약 유형

##### [ 7가지 활동 지표 ]

###### [# 2-1. 활동 이용 정도]

- 장기 이용 지표

###### [# 2-2. 활동 측면]

- 활동 다양성, 종교, 차분&활발&지적

###### [# 2-3. 활동 계절]

- 여름, 가을

##### [ 4가지 이용 지표 ]

###### [# 3-1. 효돌 동작]

- 다양한 동작 시도

###### [# 3-2. 이용 계절]

- 봄 & 여름 & 겨울형

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성

#### 17가지의 행동 패턴/생활 양식 지표

##### [ 6가지 종합 지표 ]

###### [기상/취침]

- 새벽 & 올빼미형

###### [규칙적 생활]

- 규칙적 아침 & 점심/저녁형

###### [약 복용]

- 식전 & 식후 약 유형

##### [ 7가지 활동 지표 ]

위의 17개 (6+7+4개)의 지표를 얻어내기 위해,

###### [활동 이용 정도]

- 장기 이용 지표

#### 요인 분석 (Factor Analysis) 진행

###### [활동 측면]

- 조건 1) 분산 설명 비율이 60%가 넘어

###### [활동 계절]

- 조건 2) 생성된 요인 변수가 해석 가능

- 여름, 가을

##### [ 4가지 이용 지표 ]

###### [효돌 동작]

- 다양한 동작 시도

###### [이용 계절]

- 봄 & 여름 & 겨울형

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성 - 6가지 종합 지표

변수 가공 과정 : 불필요한 변수 제거 & 변수 묶음 & 변수 추가

1

인형/유저 종합 정보

( 기존 49개 -> 가공 28개 )

성별, 연령, 산책/환기 여부, 복용 약물, 질병 종류, 기상 시간, 식사 시간 등...

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성 - 6가지 종합 지표

변수 가공 과정 : 불필요한 변수 제거 & 변수 묶음 & 변수 추가

1

인형/유저 종합 정보

( 기존 49개 -> 가공 28개 )

성별, 연령, 산책/환기 여부, 복용 약물, 질병 종류, 기상 시간, 식사 시간 등...



요인 분석 후의, HIGHEST & LOWEST 계수

	1) 새벽형	2) 올빼미형	3) 규칙적 아침식사	4) 규칙적 저녁식사	5) 식전 약 복용	6) 식후 약 복용
(+)	새벽 기상 / 아침식사 ( 1.05, 0.23 )	늦은 취침 ( 0.87 ) & 아침 기상 ( 0.11 )	정시 아침식사 ( 1.11 ), 아침 기상 ( 0.18 )	정시 저녁 & 점심 식사 ( 0.92 , 0.17 )	저녁 / 점심 / 아침 식전약 ( 0.71, 0.67, 0.62 )	아침/저녁/점심 식후약 ( 0.75, 0.70, 0.44 )
(-)	아침 기상 / 아침식사 ( -0.99, -0.09 )	이른 취침 ( -0.76 ) & 정시 점심식사 ( -0.20 )	새벽 & 점심 아침식사 ( -0.69, -0.37 )	심야 저녁식사 ( -0.91 ), 이른 취침 ( -0.09 )	아침 / 저녁 식후약 ( -0.39, -0.28 )	아침/점심/저녁 식전약 ( -0.12, -0.08, -0.06 )

[# 1-1. 수면 패턴]

[# 1-2. 규칙적 생활]

[# 1-3. 약 복용]

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성 - 7가지 활동 지표

변수 가공 과정 : 불필요한 변수 제거 & 변수 묶음 & 변수 추가

2

귀 프로그램 수행 로그

( 기존 11개 -> 가공 23개 )

7가지 활동에 대한 경험 여부, 상대적 선호도 /

계절 별 이용 월 수, 상대적 선호도 / 프로그램 이용 기간

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성 - 7가지 활동 지표

변수 가공 과정 : 불필요한 변수 제거 & 변수 묶음 & 변수 추가

2

귀 프로그램 수행 로그

( 기존 11개 -> 가공 23개 )

7가지 활동에 대한 경험 여부, 상대적 선호도 /  
계절 별 이용 월 수, 상대적 선호도 / 프로그램 이용 기간



요인 분석 후의, HIGHEST & LOWEST 계수

	1) 장기 이용	2) 여름, 가을	3) 활동 다양성	4) 종교 활동
(+)	활동 일수 ( 0.89 )	가을 / 여름 월수 ( 0.78, 0.62 )	이야기 여부 ~ 음악 일수 ( 0.86~0.26 )	종교 횟수 & 여부 ( 1.04, 0.66 )
(-)	여름 월수 ( -0.24 )	봄 횟수 ( -0.27 )	겨울 횟수 ( -0.21 )	음악 & 회상 횟수 ( -0.21, -0.19 )

[# 2-1. 활동 이용 정도]

[# 2-2. 활동 계절]

[# 2-3. 활동 측면]

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성 - 7가지 활동 지표

변수 가공 과정 : 불필요한 변수 제거 & 변수 묶음 & 변수 추가

2

귀 프로그램 수행 로그

( 기존 11개 -> 가공 23개 )

7가지 활동에 대한 경험 여부, 상대적 선호도 /

계절 별 이용 월 수, 상대적 선호도 / 프로그램 이용 기간



요인 분석 후의, HIGHEST & LOWEST 계수

6) 차분 취미	6) 활발 취미	7) 지적 취미
회상 횟수 & 여부 ( 0.90, 0.71 ) 이야기 여부 & 횟수 ( 0.16, 0.08 )	체조 횟수 & 여부 ( 0.67, 0.61 ), 퀴즈 횟수 & 여부 ( 0.37, 0.34 )	영어 횟수 & 여부 ( 0.84, 0.67 ), 이야기 여부 ( 0.08 ), 퀴즈 횟수 & 여부 ( 0.07, 0.07 )
음악 횟수 & 여부 ( -0.32, -0.23 )	음악 횟수 ( -0.41 ), 회상 횟수 ( -0.28 ), 종교 횟수 ( -0.24 )	체육 횟수 & 여부 ( -0.23, -0.17 ), 음악 횟수 & 여부 ( -0.17, -0.12 )

[# 2-3. 활동 측면]



## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성 - 4가지 이용 지표

6가지 동작 : 머리 쓰다듬 / 손 버튼 / 등 두드림 / 인체 감지 / 체조 / 퀴즈

3

인형 접속 로그

( 기존 6개 -> 가공 21개 )

이용 기간1,2 (총 접속 일수, 최초~최후 접속),  
20일 이상 사용 경험 여부, 월별 이용 여부, 6가지 동작(\*) 사용 횟수

## 3-1. 고객 군집화

### (1) 군집화를 위한 변수 생성 - 4가지 이용 지표

6가지 동작 : 머리 쓰다듬 / 손 버튼 / 등 두드림 / 인체 감지 / 체조 / 퀴즈

3

인형 접속 로그

( 기존 6개 -> 가공 21개 )

이용 기간1,2 (총 접속 일수, 최초~최후 접속),

20일 이상 사용 경험 여부, 월별 이용 여부, 6가지 동작(\*) 사용 횟수



요인 분석 후의, HIGHEST & LOWEST 계수

	1) 다양한 동작 시도	2) 봄 이용	3) 여름 이용	4) 겨울 이용
(+)	퀴즈 ~ 머리 쓰다 ( 0.90~0.60 )	4,5,3월 이용 ( 0.83 ~ 0.69 )	7,6,8월 이용 ( 0.84~0.67 )	11,12,10,1,2월 이용 ( 0.78 ~ 0.53 )
(-)	20일 이상 이용 ( -0.12 )	9,10,8월 여부 ( -0.27~ -0.14 )	1월 이용 ( -0.13 )	6,4,5,7월 이용 ( -0.16~ -0.11 )

[# 3-1. 효율 동작]

[# 3-2. 이용 계절]

## 3-1. 고객 군집화

### (2) 군집화 알고리즘



데이터 (군집화 대상)

17가지의 행동 패턴/생활 양식 지표 를 가진 1729개의 인형

군집화 알고리즘

SOM (Self-Organizing Map)

알고리즘 선정 이유

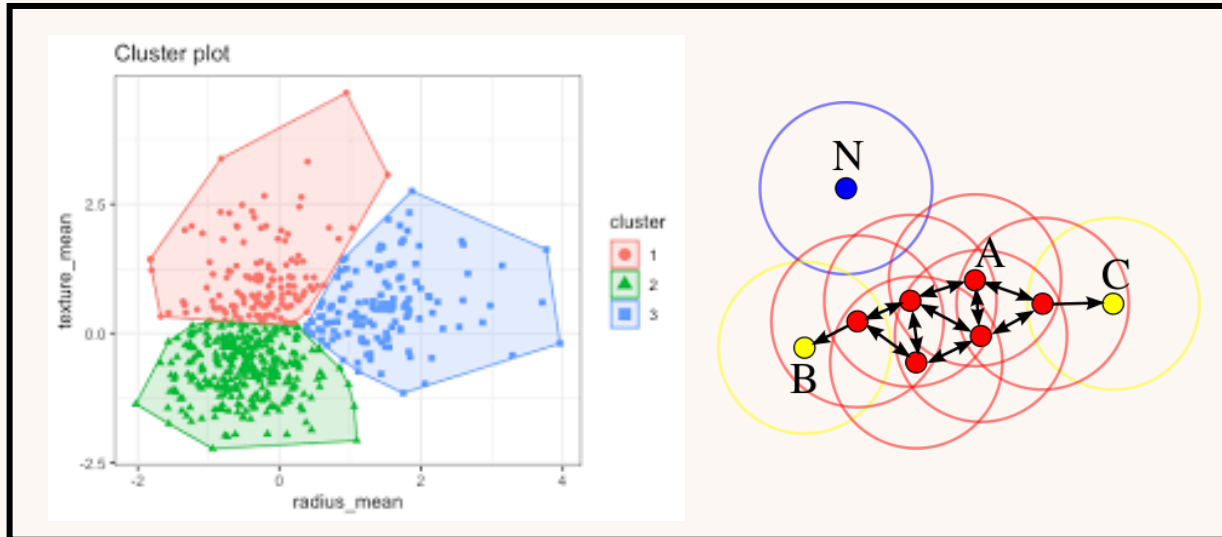
(축을 사용하여) 군집 들 간의 상대적 거리 비교 가능

## 3-1. 고객 군집화

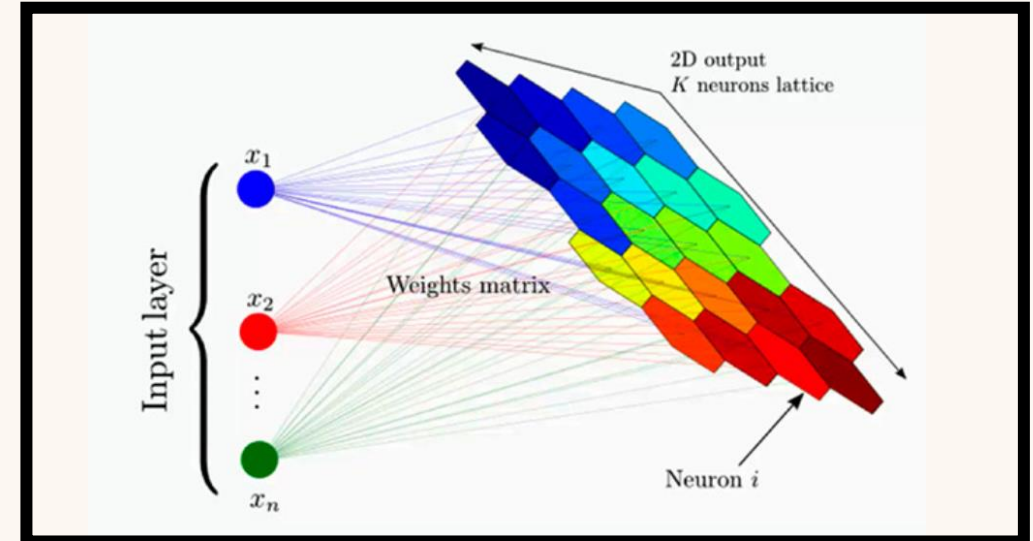
### (2) 군집화 알고리즘

군집화 알고리즘

SOM (Self-Organizing Map)



기존 알고리즘들 : 군집 간 순서/비교 용이 X



SOM : 격자 형식의 비교가능한 군집

## 3-1. 고객 군집화

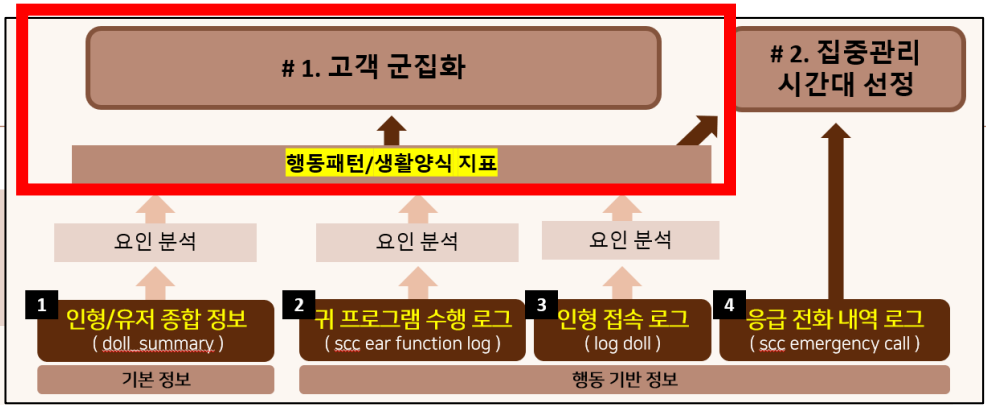
### (2) 군집화 알고리즘

군집화 알고리즘

SOM (Self-Organizing Map)

Cluster # 1, 2, .., 8, 9

기존 알고리즘들 : 군집 간 순서/비교 용이 X



Cluster # (1,1) ~ (3,3)

SOM : 격자 형식의 비교가능한 군집

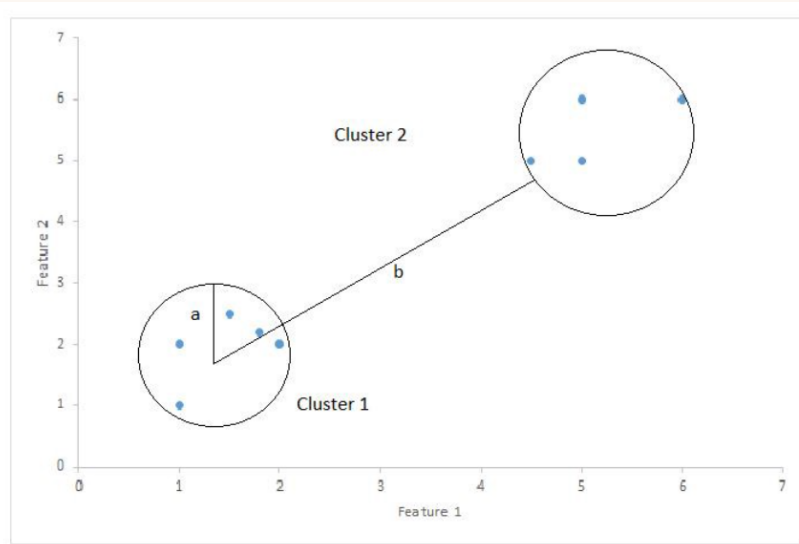
## 3-1. 고객 군집화

### (2) 군집화 알고리즘

군집화 알고리즘

SOM (Self-Organizing Map)

Algorithm Comparison ( with Silhouette score )

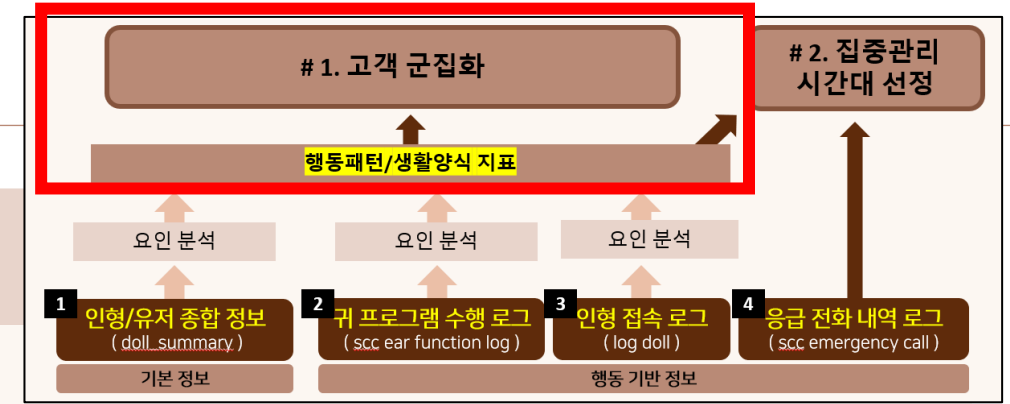


$$\text{Silhouette Score} = (b-a)/\max(a,b)$$

a = "군집 **내**" 평균 거리

b = "군집 **간**" 평균 거리

-1~1 사이의 값을 가지는, "**클수록**" 좋은 스코어



## 3-1. 고객 군집화

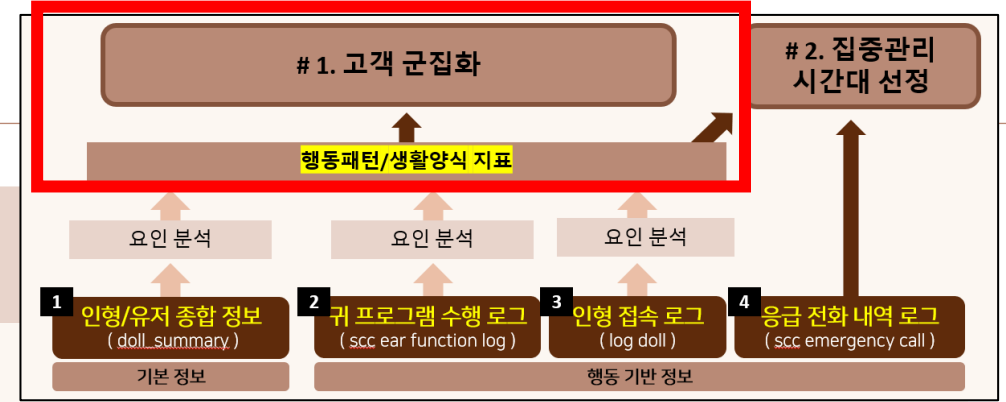
### (2) 군집화 알고리즘

군집화 알고리즘

SOM (Self-Organizing Map)

### Algorithm Comparison ( with Silhouette score )

		Silhouette score
K-means	Elbow method 통해 최적의 K 설정	0.26
DBSCAN	Grid Search로 최적의 epsilon 탐색	0.22
Hierarchical Clustering	Tree structure 확인을 통한 적절한 cluster 개수 지정	0.18
SOM	2x2, 2x3, 3x3, 4x3을 통해 최적의 cluster 개수 지정	0.31

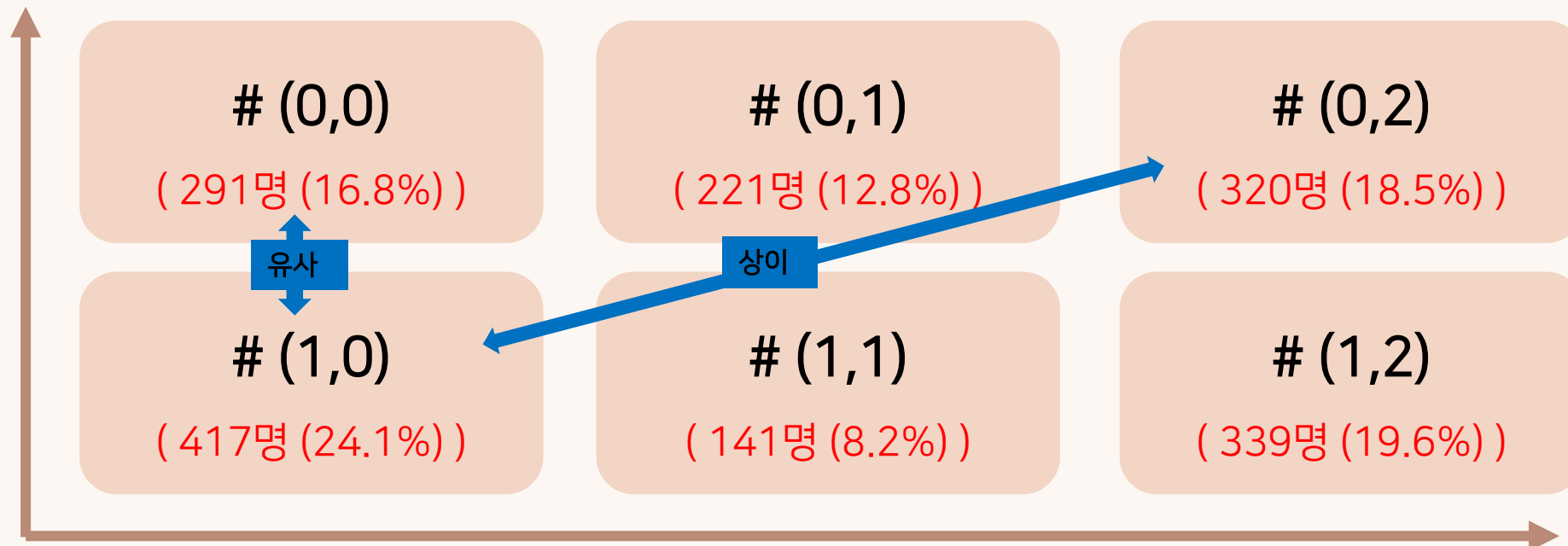


## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)

군집화 결과 : 6(= 2x3)개의 고객 군집

( 총 1,729개의 인형 )



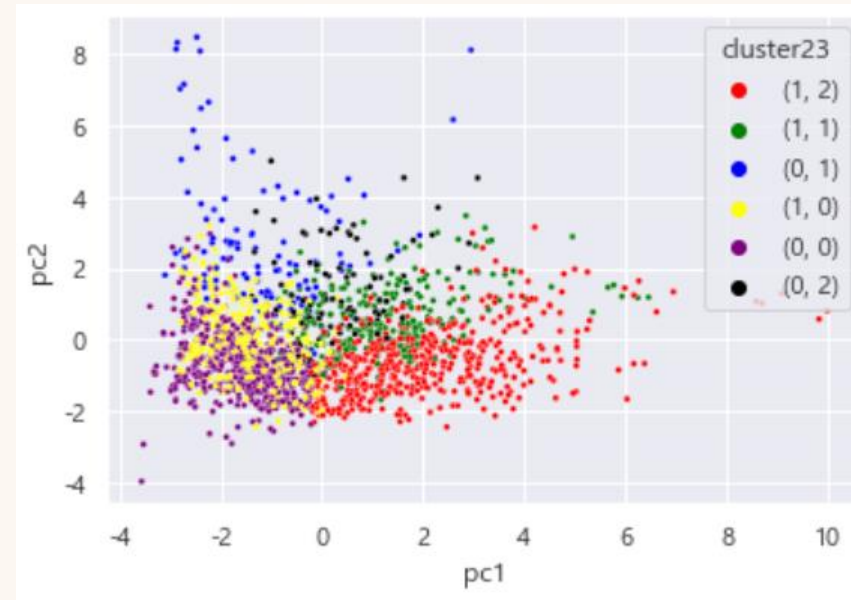
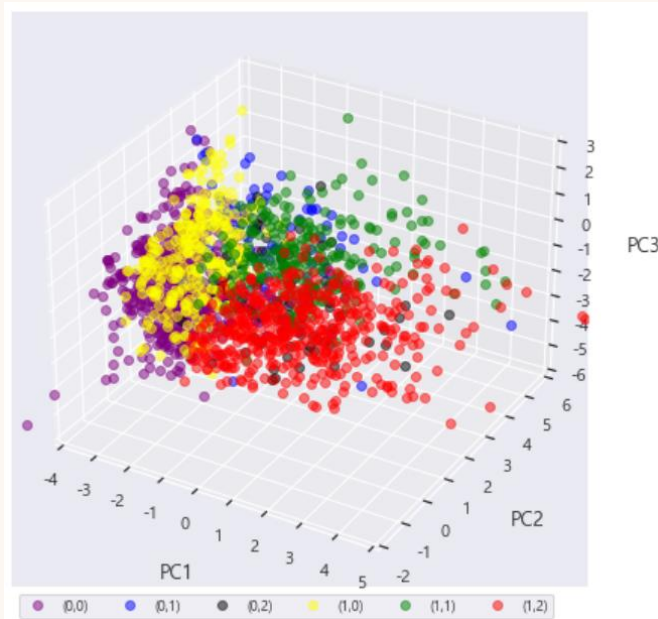


## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)

군집화 결과 : 6(= 2x3)개의 고객 군집

군집별로 17개의 행동패턴/생활양식 지표를, 주성분분석(PCA)를 사용하여 2, 3차원 시각화

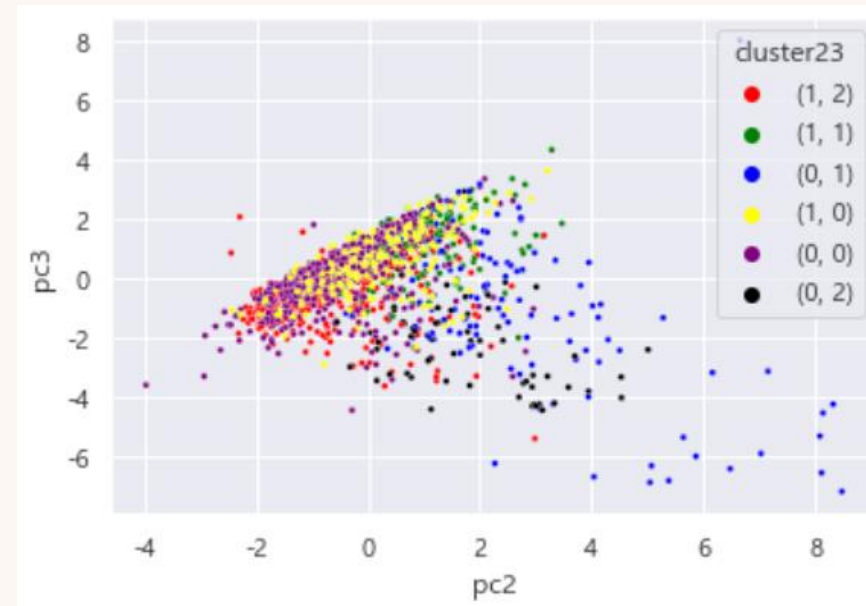
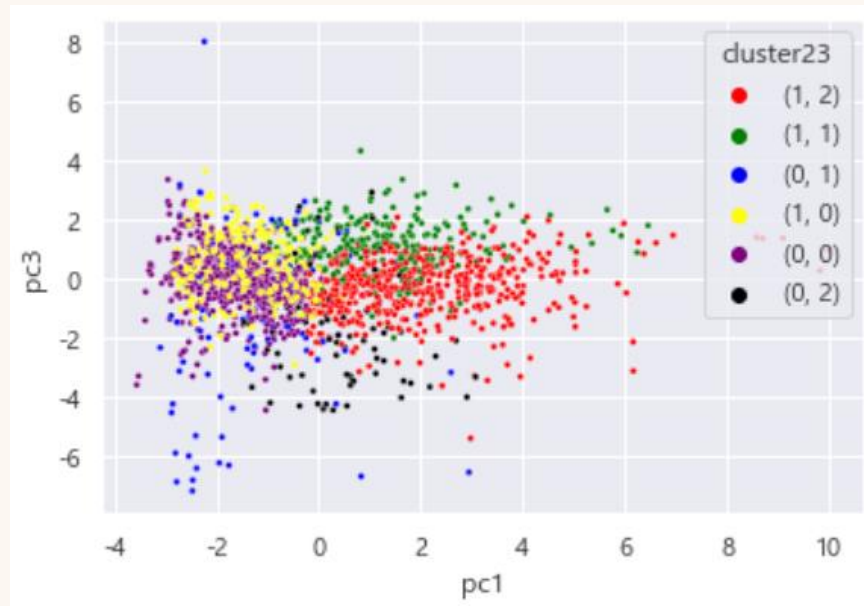


## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)

군집화 결과 : 6(= 2x3)개의 고객 군집

군집별로 17개의 행동패턴/생활양식 지표를, 주성분분석(PCA)를 사용하여 2, 3차원 시각화



## 3-1. 고객 군집화

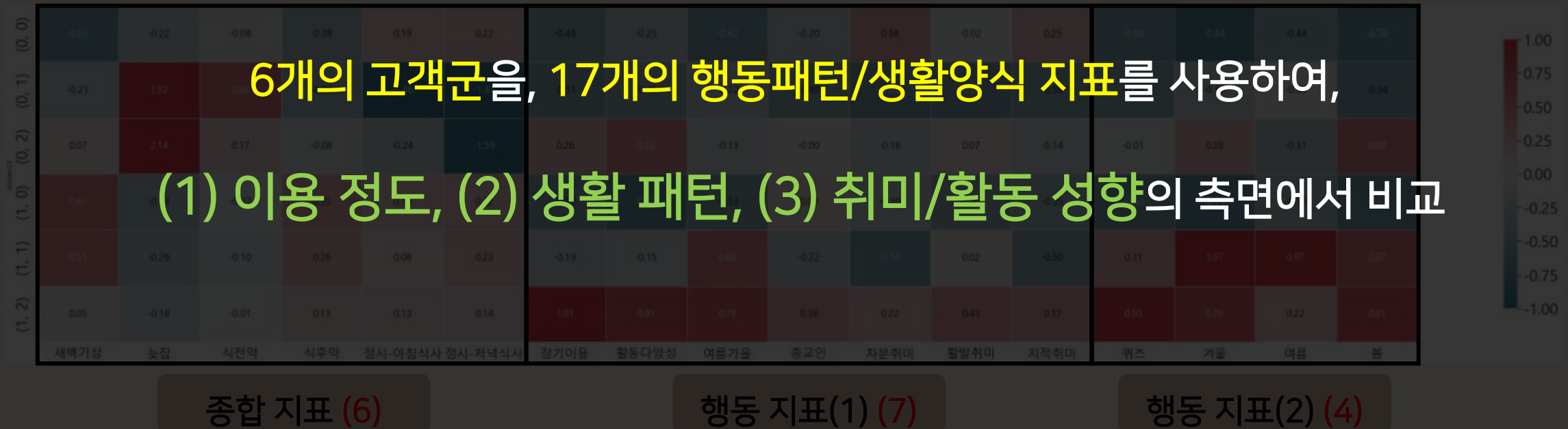
### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)



17개의 행동패턴/생활양식 지표 에서 군집 간의 비교분석

## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)



17개의 행동패턴/생활양식 지표 에서 군집 간의 비교분석

## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)

군집 별 주요 특징 - (1) 이용 정도

(-) 요인

(+) 요인



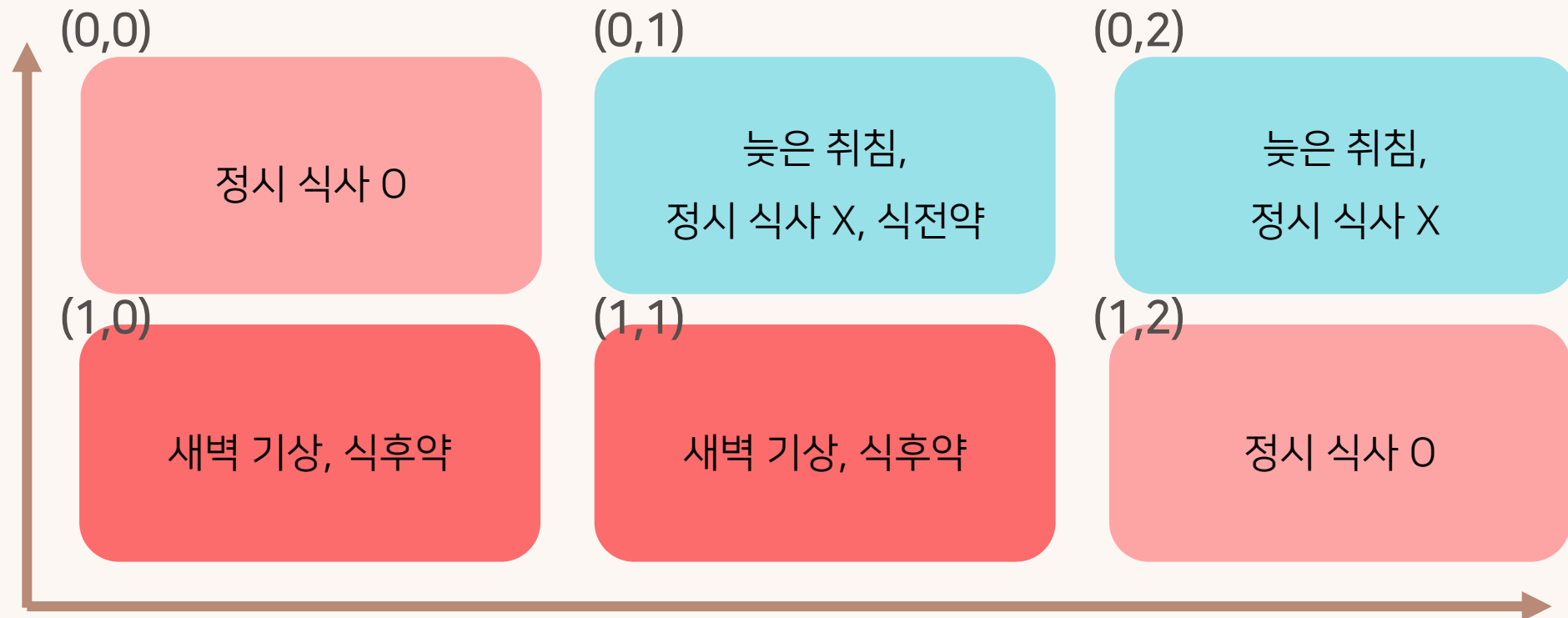
## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)

군집 별 주요 특징 - (2) 생활 패턴 (기상, 식사, 약)

(-) 요인

(+) 요인



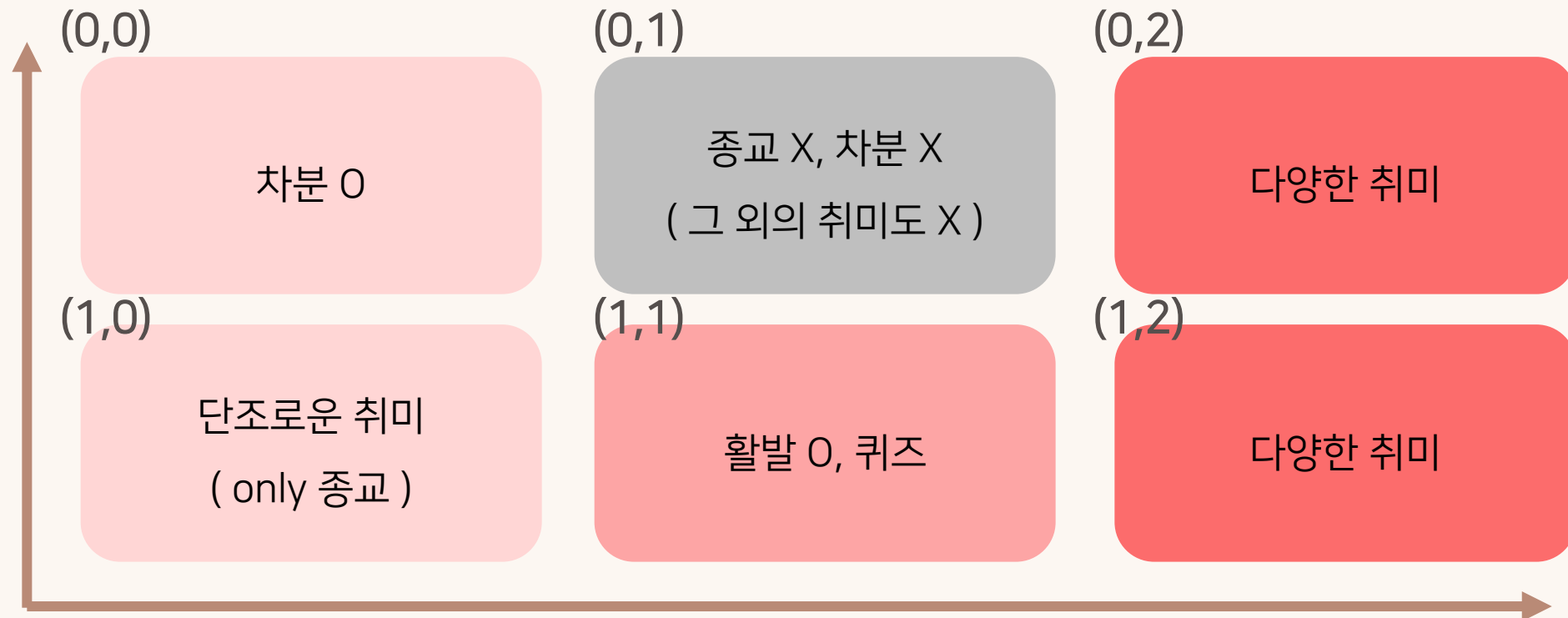
## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)

군집 별 주요 특징 - (3) 취미/활동 성향 - 활동 다양성 측면

단조

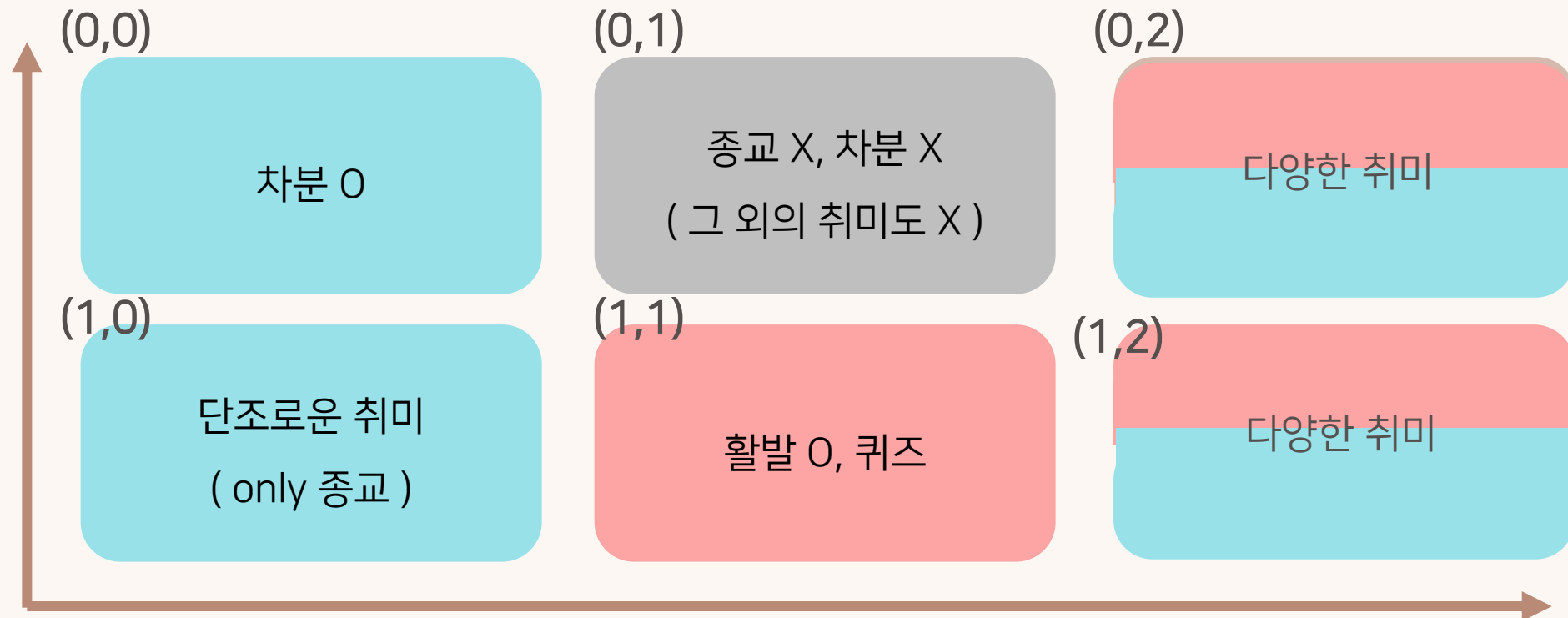
다양



## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)

군집 별 주요 특징 - (3) 취미/활동 성향 - 활동 성향 측면





## 3-1. 고객 군집화

### (3) 군집 별 페르소나 (+ 행동/생활 패턴 비교 분석)

군집 별 주요 특징 - SUMMARY- 페르소나

취미/활동 생활패턴 이용정도

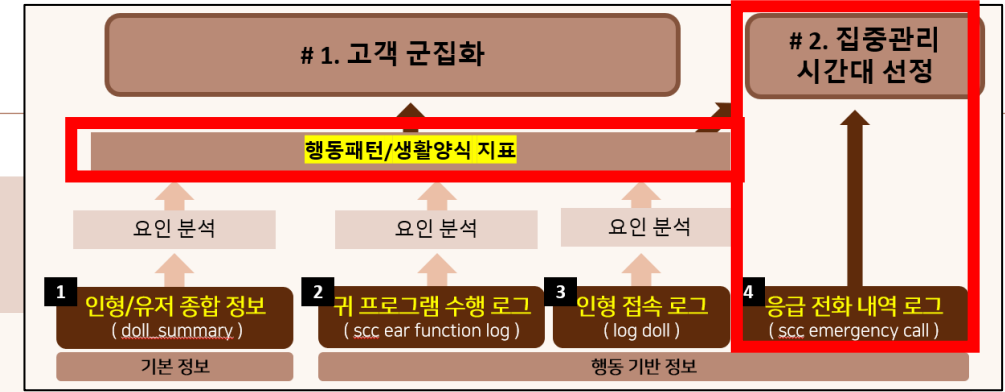


### 3. 고객 군집화 & 집중관리 시간대 예측

#### (2) 집중관리 시간대 예측

## 3-2. 집중 관리 시간대 예측

### (1) 응급 전화 내역 기반 예측



데이터

응급 전화 내역 로그 (연/월/일/시) ( + 군집 인덱스 정보 )

예측 대상

특정한 월 &amp; 시간대(\*)에 응급 전화를 걸었는지 여부

사용 알고리즘

Multi-task Neural Network

알고리즘 선정 이유

SIMPLE + EFFICIENT + REASONABLE

\* [ 시간대 선정 기준 ] 새벽 : 03~06 / 아침 : 07~11 / 이른 오후 : 12~16 / 저녁 : 17~20 / 밤 : 21~02

## 3-2. 집중 관리 시간대 예측

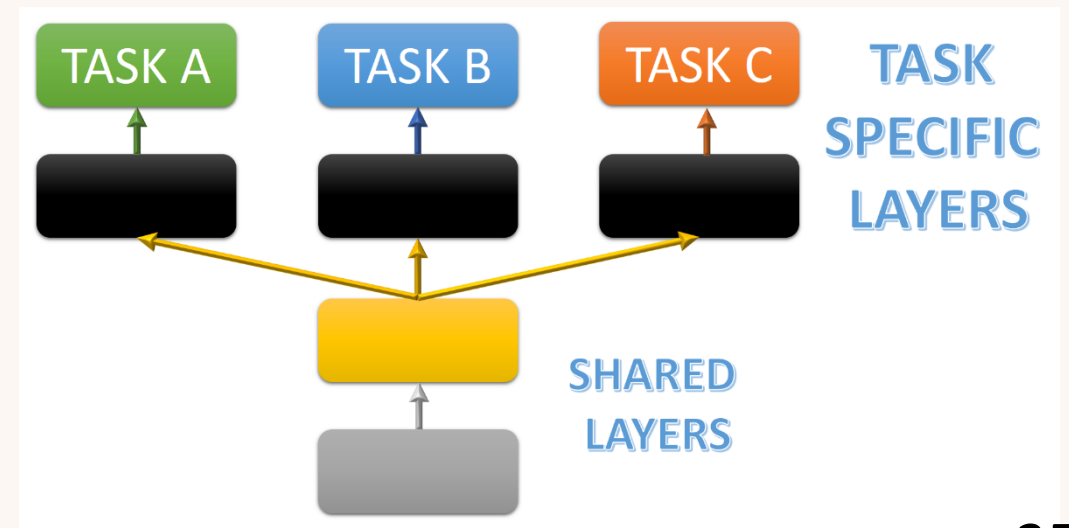
### (1) 응급 전화 내역 기반 예측

## Multi-task Learning

[ 소개 ] 여러 개의 task를 하나의 모델 만을 사용해서 푸는 알고리즘

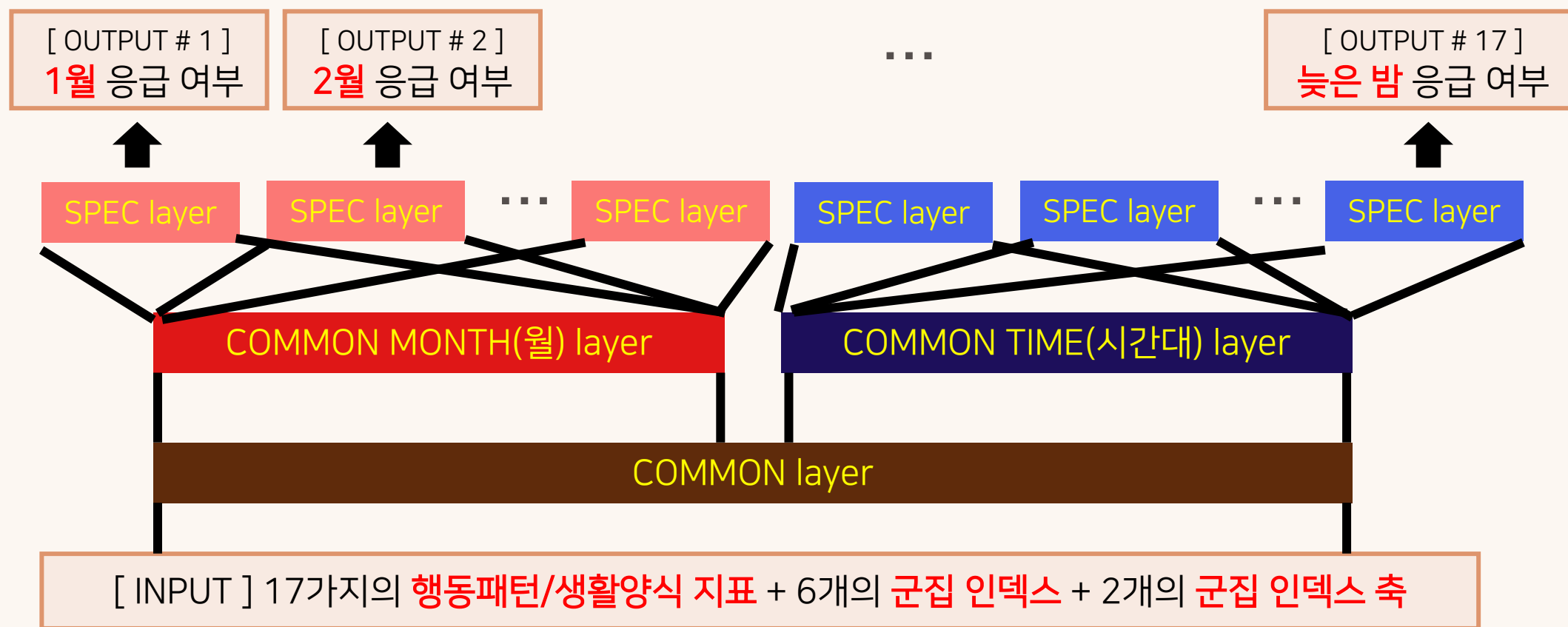
[ 가정 ] task들 간의 상관관계가 존재할 것이다

[ 구성 ] Shared layers & Task Specific Layers



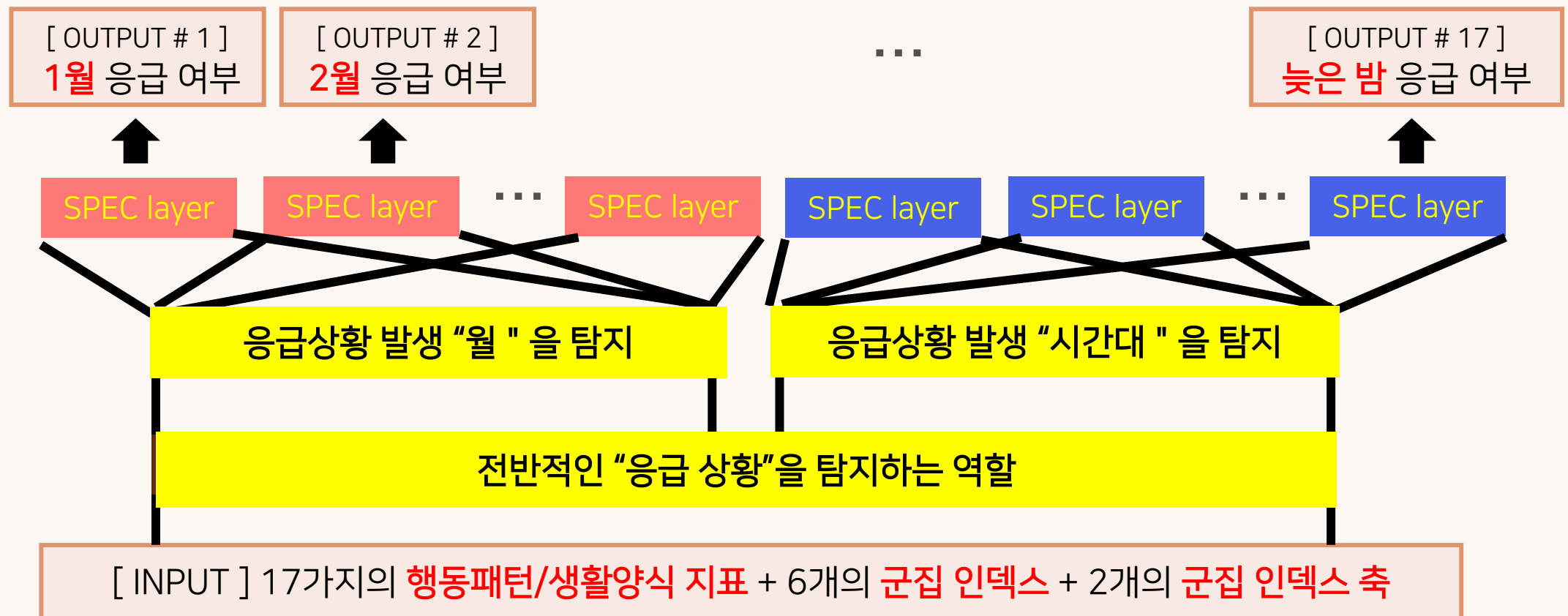
## 3-2. 집중 관리 시간대 예측

### (2) 계절/시간대 예측 알고리즘 ( Multi-task Learning )



## 3-2. 집중 관리 시간대 예측

### (2) 계절/시간대 예측 알고리즘 ( Multi-task Learning )



## 3-2. 집중 관리 시간대 예측

### (2) 계절/시간대 예측 알고리즘 ( Multi-task Learning )

알고리즘	예측 성능 ( 17개 반응변수 모델의 평균 )			모델 개수	모델 학습 시간
	F1	ACC	AUROC		
Logistic Regression	0.571	0.780	0.838	17	22.8 sec
Random Forest	0.897	0.952	0.985	17	447.9 sec
Multi-task NN	<b>0.940</b>	<b>0.969</b>	<b>0.988</b>	<b>1</b>	<b>5.1 sec</b>

[ 공통 사항 ]

(1) **Logistic Regression & RF** : 5-fold Cross Validation / **Multi-task NN** : 고정된 validation data 20%

(2) **모델 학습 시간** = Multi-task NN 외의 모델의 경우, 17개의 모델을 학습하는데 걸린 시간

(3) **분류 기준 (threshold)** : 1/0 분류 기준을 0.1,0.2,..0.9로 시도 후, F1-score가 가장 높은 분류기준을 선택

## 3-2. 집중 관리 시간대 예측

### (2) 계절/시간대 예측 알고리즘 ( Multi-task Learning )

#### 모델 1) Logistic Regression

```
best_cols
✓ 0.6s
['thres0.3_emergency_month1',
 'thres0.2_emergency_month2',
 'thres0.2_emergency_month3',
 'thres0.2_emergency_month4',
 'thres0.2_emergency_month5',
```

Ex) 1월 응급여부 예측 모델의 최적 분류 기준 = 0.3

#### 모델 2) Random Forest

Random Search를 통해, 17개 모델 별로 각각의 최적의 하이퍼파라미터를 사용

```
best_cols
✓ 0.3s
['thres0.4_emergency_month1',
 'thres0.3_emergency_month2',
 'thres0.4_emergency_month3',
 'thres0.4_emergency_month4',
```

```
[ y ]emergency_month1
best_params : {'n_estimators': 196, 'min_samples_split': 19, 'max_depth': 15}
[ y ]emergency_month2
best_params : {'n_estimators': 178, 'min_samples_split': 18, 'max_depth': 11}
```

17개 모델 별로 각각 상이한 최적의 구조 (하이퍼 파라미터) 적용



## 3-2. 집중 관리 시간대 예측

### (2) 계절/시간대 예측 알고리즘 ( Multi-task Learning )

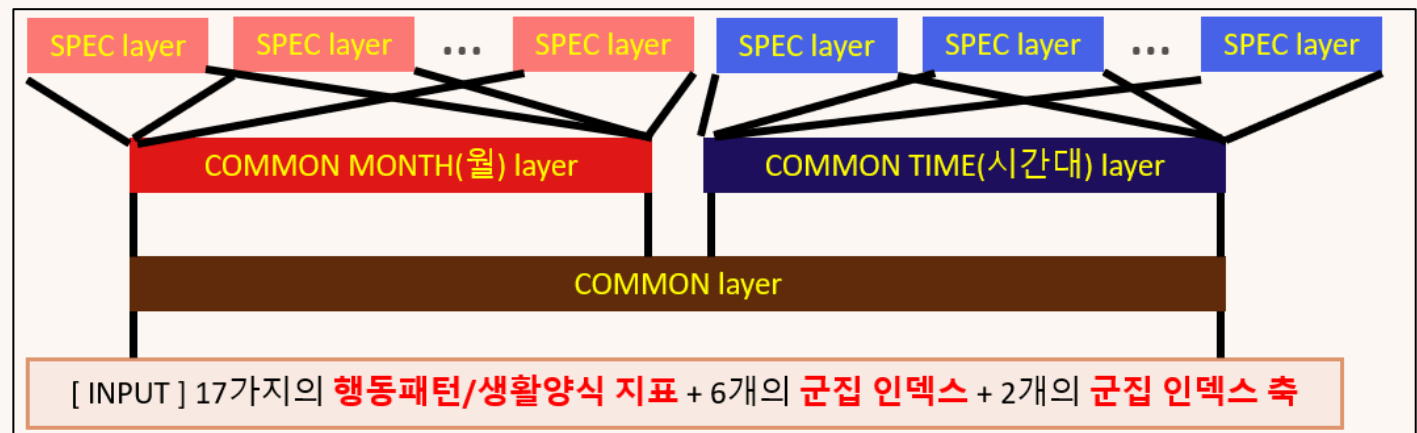
#### 모델 3) Multi-task NN

##### [1] Structure

- Common layer 1층 ( 64개의 unit )
- Common MONTH layer 1층 ( 32개의 unit ) & Common TIME layer 1층 ( 32개의 unit )
- Task-specific layer ( 8개의 unit )

[2] Epoch : 10

[3] Optimizer : Adam



## 3-2. 집중 관리 시간대 예측

### (2) 계절/시간대 예측 알고리즘 ( Multi-task Learning )

#### 모델 3) Multi-task NN

##### 1) Efficient & Fast

- RF & Log Reg : 17개의 모델 필요 vs Multi-task NN : 1개의 모델 필요
- 학습 속도 : Multi-task NN (5.1 초) < Log Reg (22.8 초) < RF (447.9 초)

##### 2) Accurate

- F1 score, Accuracy, AUROC 지표에서, 다른 알고리즘들 모두 압도

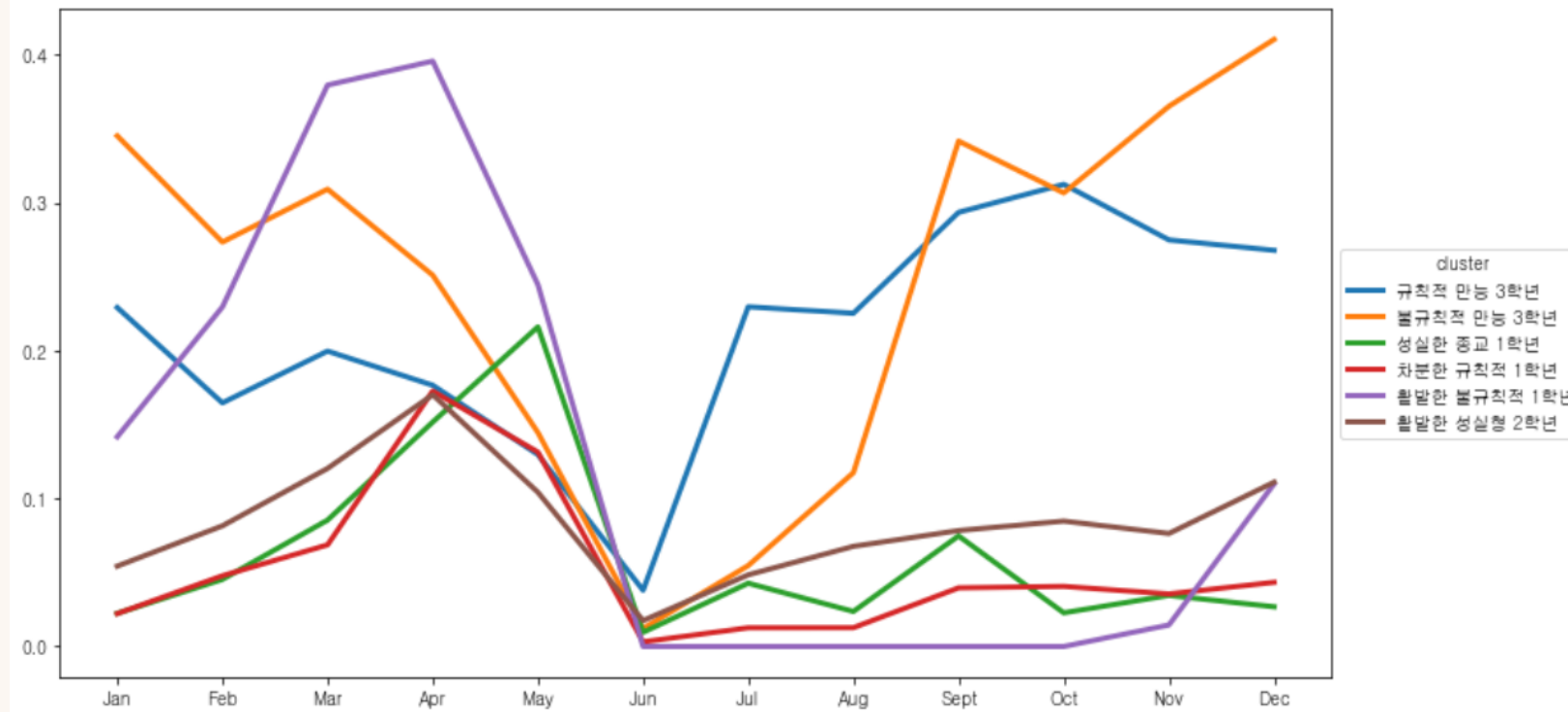
##### 3) Reasonable

- 상관성이 높은 여러 반응 변수들을, "task 공동" & "task 개별" 파라미터로 나뉜 모델 아키텍처

## 3-2. 집중 관리 시간대 예측

### (3) 군집 별 집중 관리 계절/시간대

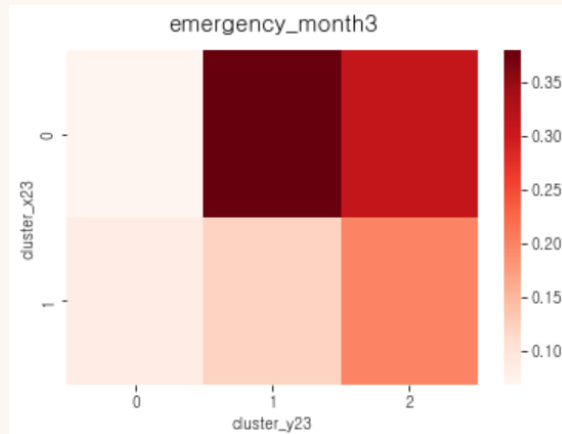
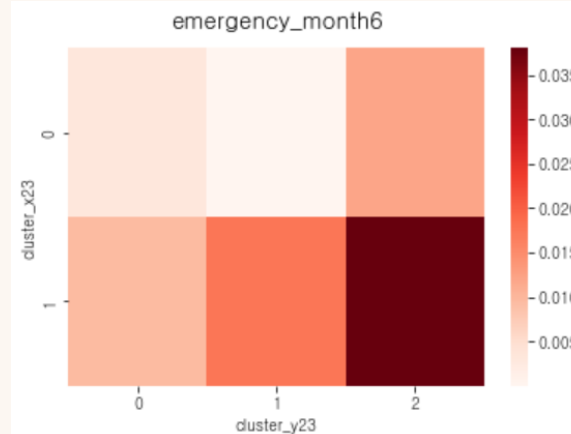
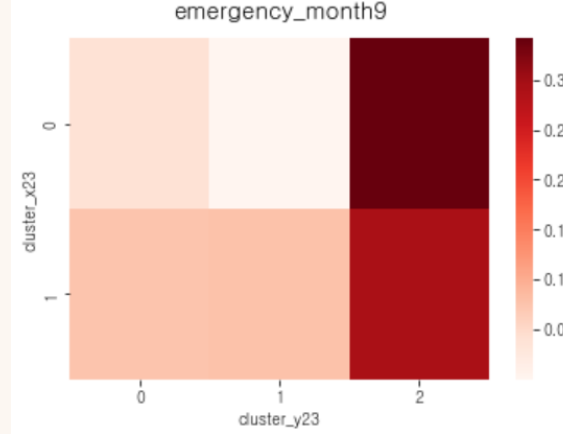
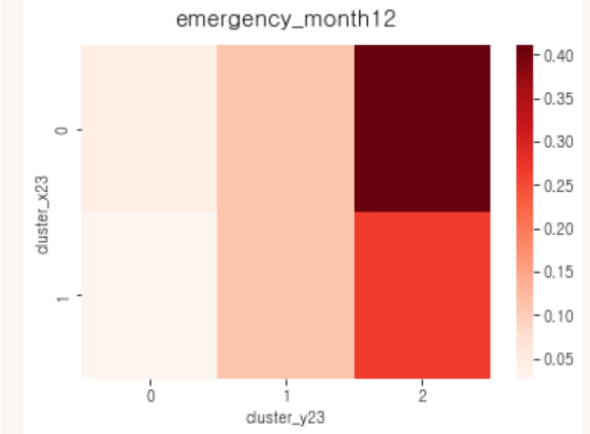
월별



군집 별로, 응급 상황 발생 여부 확률이 **"시간대 별"** & **"월 별"**로 차이를 보인다

## 3-2. 집중 관리 시간대 예측

### (3) 군집 별 집중 관리 계절/시간대

**월별****3월****6월****9월****12월**

모든 시각화는 **효린이** 통해 상세히 확인 가능

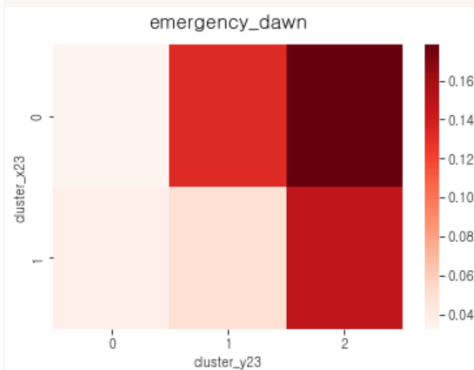
군집 별로, 응급 상황 발생 여부 확률이 **"시간대 별" & "월 별"**로 차이를 보인다

## 3-2. 집중 관리 시간대 예측

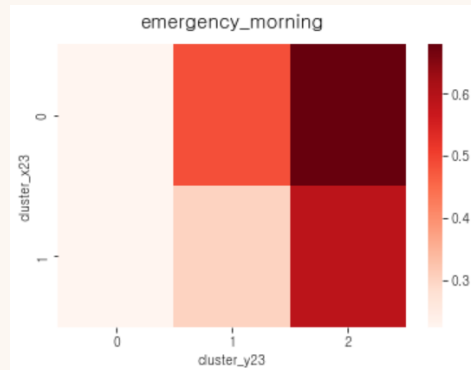
### (3) 군집 별 집중 관리 계절/시간대

시간대별

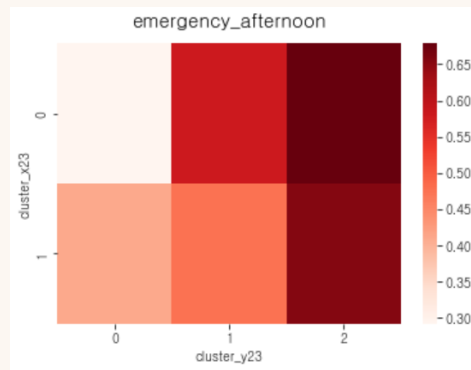
새벽



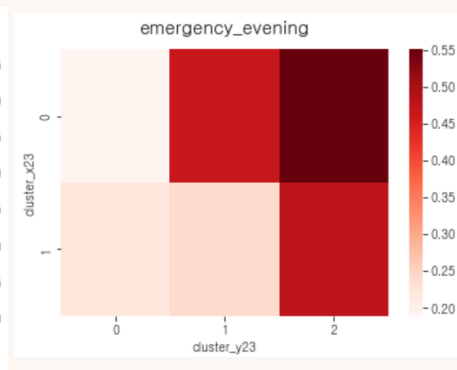
아침



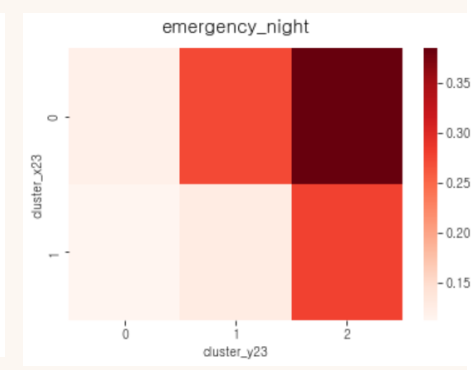
점심



저녁



밤

모든 시각화는 **효린이** 통해 상세히 확인 가능군집 별로, 응급 상황 발생 여부 확률이 **"시간대 별" & "월 별"**로 차이를 보인다

## 4. 군집 별 제언

## 4. 군집 별 요약 + 제언

### SUMMARY

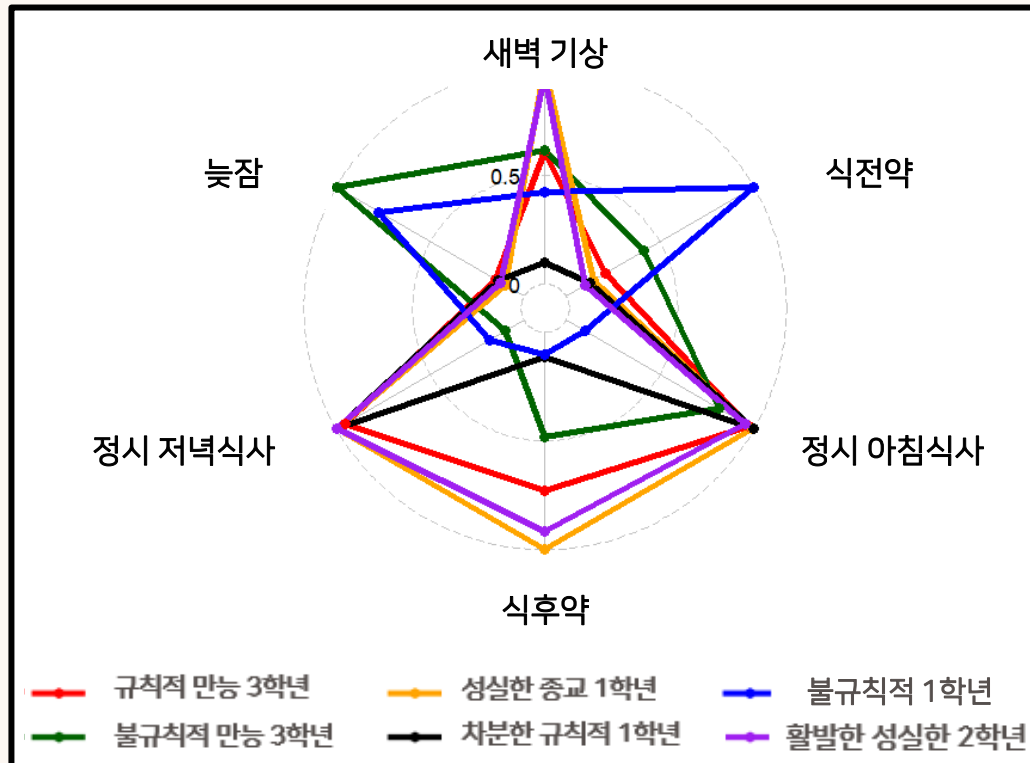
[요약] Radar Chart로 **종합 + 활동 + 이용** 지표 비교분석 ( feat. **효린이** )

[제언] 계절 별 / 시간대 별 **관리 시간대** 선정

## 4. 군집 별 요약 + 제언

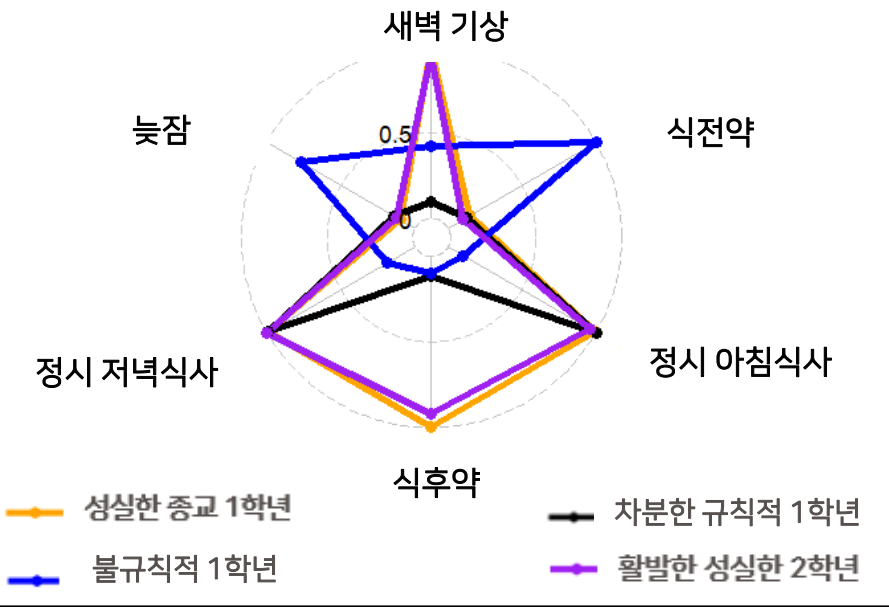
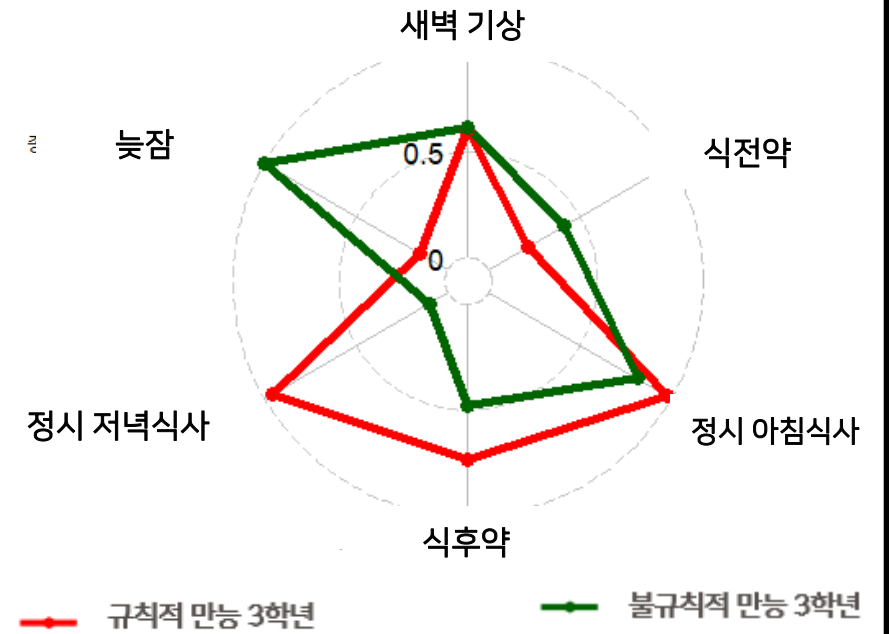
### 종합 + 활동 + 이용 지표 비교분석

#### 종합 지표 6가지



장기(3학년)

단.중기(1,2학년)

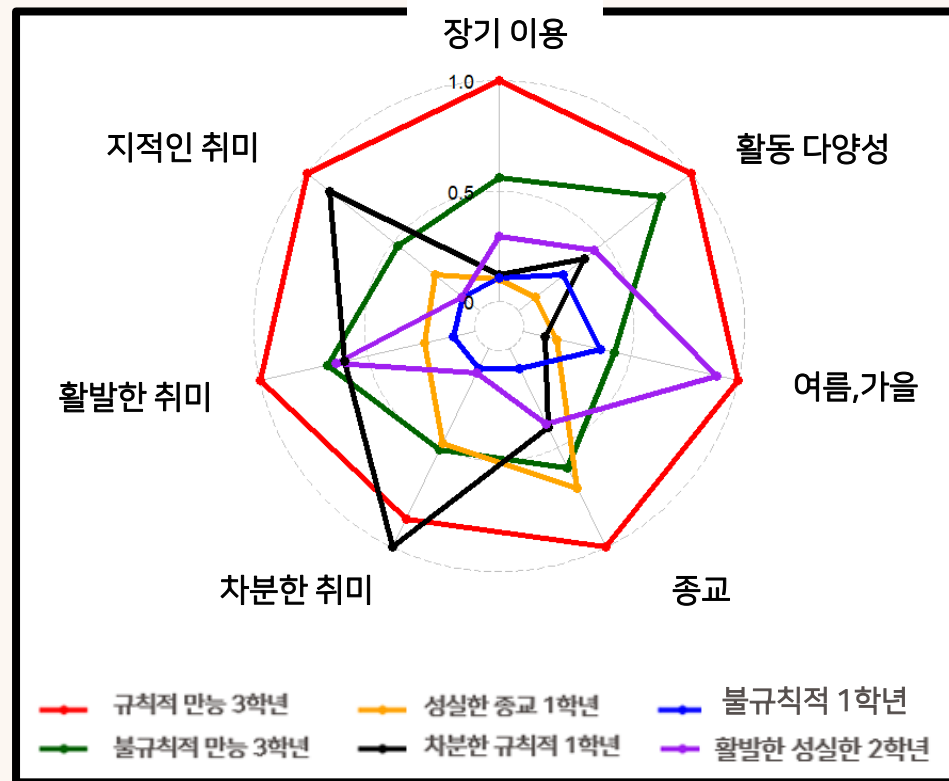




## 4. 군집 별 요약 + 제언

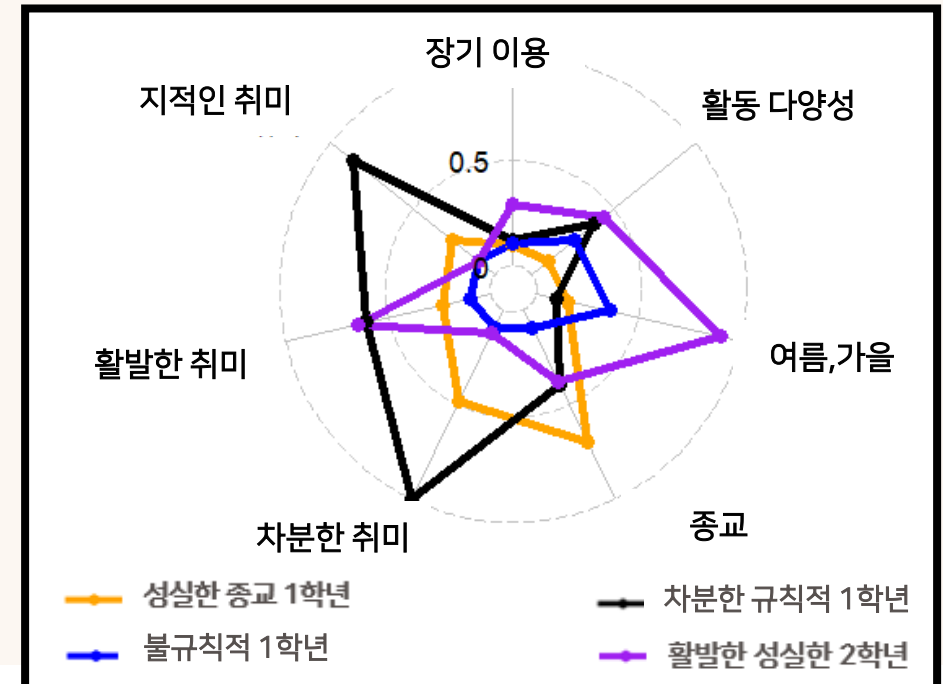
### 종합 + 활동 + 이용 지표 비교분석

#### 활동 지표 7가지



장기(3학년)

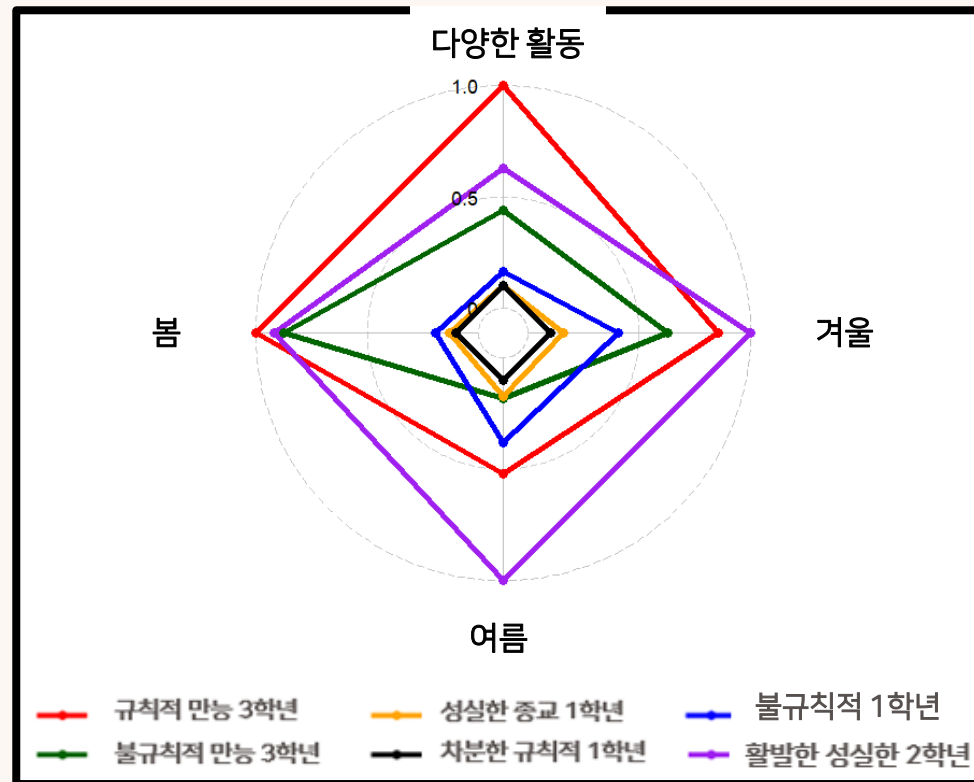
단.중기(1,2학년)



## 4. 군집 별 요약 + 제언

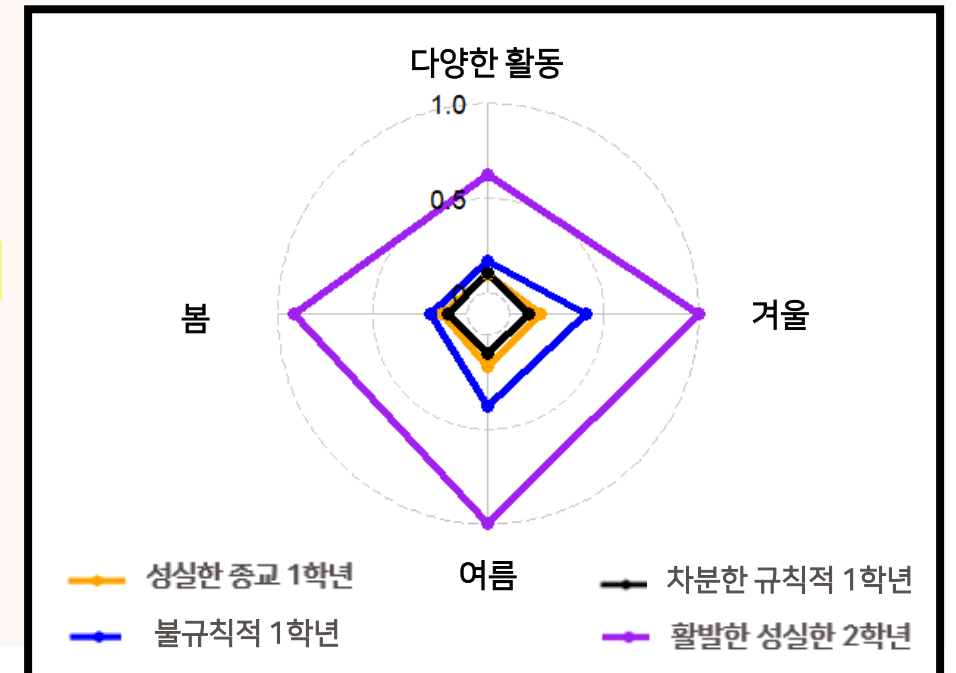
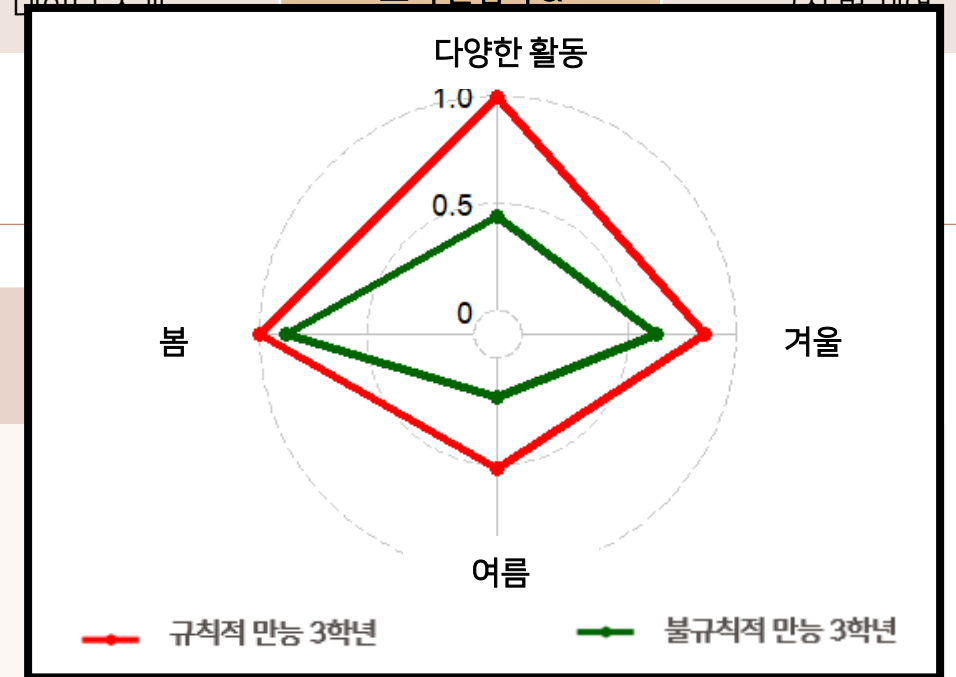
### 종합 + 활동 + 이용 지표 비교분석

#### 이용 지표 4가지



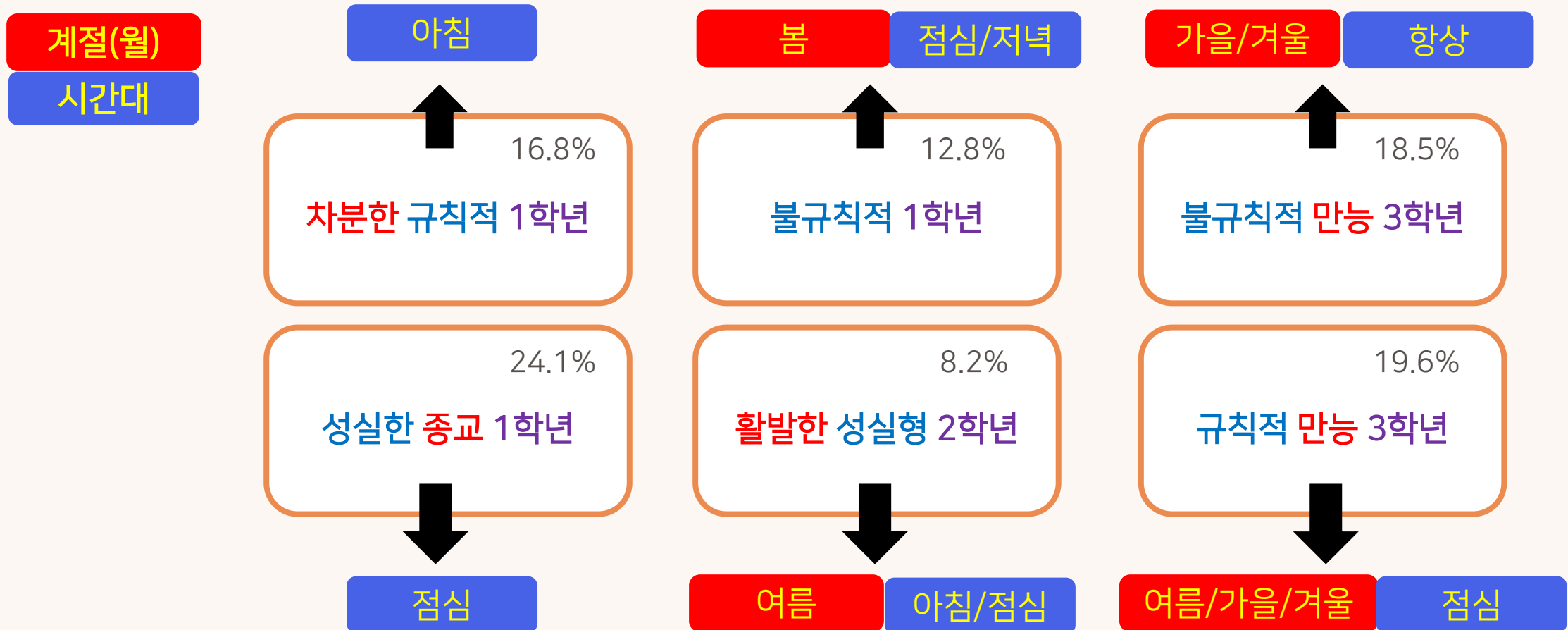
장기(3학년)

단.중기(1,2학년)



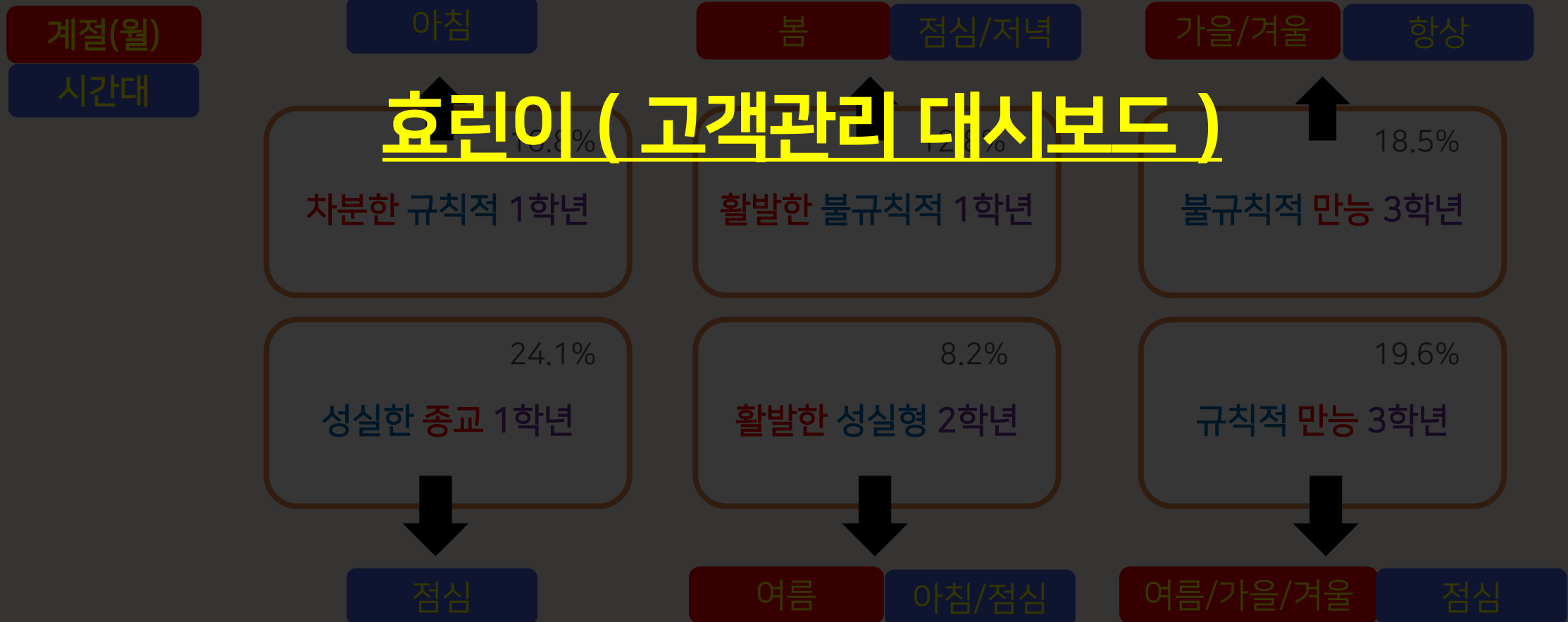
## 4. 군집 별 요약 + 제언

### 계절별 / 시간대별 관리 시간대 선정



## 4. 군집 별 요약 + 제언

계절별 / 시간대별 관리 시간대 선정



## 4. 군집 별 요약 + 제언

### 군집 별 제언

차분한 규칙적 1학년

성실한 종교 1학년

불규칙적 1학년

활발한 성실형 2학년

불규칙적 만능 3학년

규칙적 만능 3학년

단기 이용

중기 이용

장기 이용

정시 O 식사

정시 X 식사

정시 X 식사

정시 O 식사

새벽 기상

늦은 취침

새벽 기상

늦은 취침

차분 O 취미

종교 취미

활발 O 취미

만능 취미

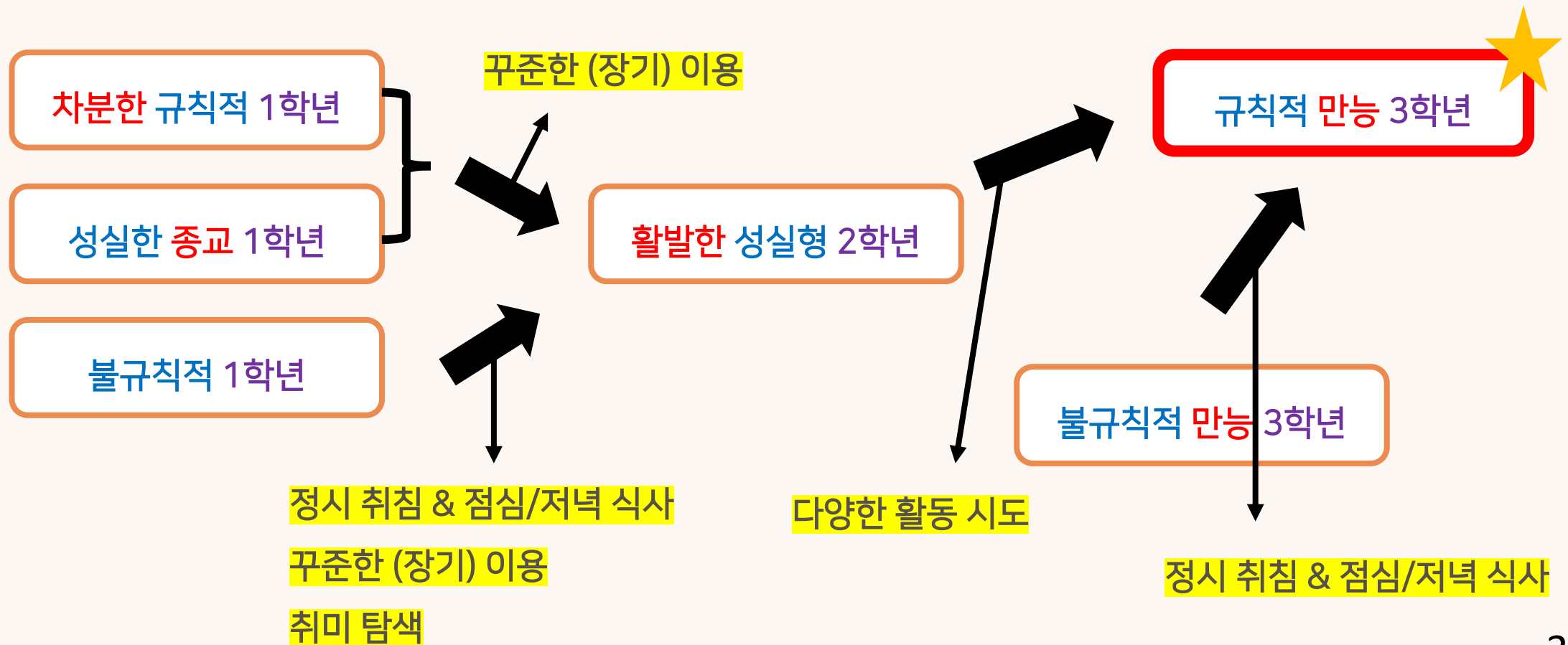
식후약

식전약

식후약

## 4. 군집 별 요약 + 제언

### 군집 별 제언



# 추후 가능 분석 ( Future Works )

## # 1 ) 다양한 태스크 해결에 도움을 줄 수 있는 행동패턴/생활양식 지표

고객 군집화를 위해 사용했던 “17가지 행동패턴/생활양식 지표”를,

단지 ‘관리 시간대 예측’만이 아닌, 다양한 task를 위해서 사용될 수 있다. ( Ex) 질병 여부 예측, 설문 응답 보정 )

## # 2) 정교한 군집화 ( = 군집화를 위한 추가 변수 생성 )

유저의 로그 정보를, 축약하여 정적인 정보로 사용하는 것이 아닌,

시간적 정보를 반영한 잠재 변수를 생성하여 ( ex. Time Series Representation을 뽑아내서 ),

군집화를 위한 추가적인 변수로 사용할 수 있다.

감사합니다!