

2025 동원 x 카이스트 AI Competition

목차

1. 페르소나 설계 방식

2. 예측값 산출 시뮬레이션

3. 코드 실행 방법

4. 제품별 시장 분석

증빙자료가 필요한 부분은 [숫자] 로 표시됨

1. 페르소나 설계 방식

- 페르소나 생성에 사용한 모델
- 프롬프트 근거
- 프롬프트 구성 로직
- 페르소나 JSON 구조
- 프롬프트, 페르소나 구성 결론
- 페르소나 생성 시 사용한 프롬프트 템플릿
- 페르소나에게 부여된 속성
- 예측 파이프라인 개요

페르소나 생성에 사용한 모델

벤치마크 점수 출처 : Salesforce AI Research

모델명	정확도 (Accuracy)	속도 (Speed)	신뢰 & 안전성 (Trust & Safety)
Gemini 2.5 Pro	3.48	3.9	76
Gemini 2.5 Flash	3.45	2.4	83
Gemini 2.5 FlashLite	3.42	1.7	81
Gemini 2.0 Flash	3.36	1.8	82
Gemini Pro 1.5	3.25	5.5	91
Gemini 2.0 FlashLite	3.2	1.4	82

- 왜 Gemini 2.5 flash 모델을 사용했는가? [1]

- 대회의 핵심을 ‘얼마나 현실에 가까운 페르소나를 정교하게 생성하는가’로 설정
- 예측의 ‘신뢰도’와 ‘정교함’이 평가의 핵심 기준이 될 거라 판단
- Gemini 2.5 Flash는 Accuracy(3.45)에서 경쟁 모델인 2.0 Flash(3.36)를 확실히 앞서는 성능을 보임
- 특히 Trust & Safety(83%)에서는 모든 flash 모델 중에서 제일 높음

프롬프트 근거

모델명	설명
역할·과업 선언(헤더)	모델 역할 고정 + 임무·출력 규칙 명시
맥락 주입(컨텍스트)	[제품 정보]에 product_name 삽입 [타겟 고객 프로필]을 제품명 조건문으로 동적 주입(도메인 힌트)
스키마 고정(제약 섹션)	[속성 가이드]에 허용값(폐쇄형 레이블) 열거 → 자유도 축소·후처리 용이
출력 예시(원샷 포맷팅)	샘플 1건으로 키·타입·배열 길이(12개월) 시연 → 형식 모사 유도
매개변수화	num_personas를 프롬프트에 주입(대량 생성 일관성)

기대 효과

- **출력 안정성·환각 감소:** 섹션화된 지시 + 스키마·허용값 고정
- **현실성·일관성 향상:** 도메인 그라운딩(타겟 프로필 조건부 주입)
- **학습 부담 감소:** 원샷 예시로 포맷 빠른 학습
- **운영 안정성:** 재시도·캐시로 실패 복구 및 비용/지연 관리

프롬프트 근거

코드 규칙	논문 개념	설명
create_product_specific_prompt 헤더/섹션	Instruction Structuring	규칙·맥락·출력 분리로 오류·환각 감소
맥락 주입(컨텍스트)	Grounded Prompting	도메인 고정으로 일관성↑
스키마 고정(제약 섹션)	Constrained/Schema	파싱/후처리 안정성↑
출력 예시(원샷 포맷팅)	Few-shot(Format Imitation)	포맷 모사 유도

프롬프트 구성 로직

본 프롬프트는 단순한 질문 응답을 넘어 제품별 소비자 맥락을 정밀하게 반영한 **구조적 데이터(JSON)** 생성을 목표. 이를 위해 프롬프트는 세 가지 축으로 설계.

1. 타겟 고객 프로필 매핑

- 제품, 카테고리에 따라 핵심 소비자 집단 설정

ex: 하이그릭요거트 → 건강 및 자기 관리에 관심이 높은 20~40대 여성, 온라인 채널 구매 선호.

동원맛참 → 20~40대 1인가구(편의성과 트렌드 지향), 30~50대 주부

참치액 진/프리미엄 → 30~60대 주부, 요리 애호가, 프리미엄 제품 선호

2. 속성 정의

- 프롬프트에 포함시킬 필수 속성 명시

ex: age, gender, occupation, income_level, household_size, lifestyle, media_consumption, ...

일관된 스키마로 JSON 객체를 생성하도록 유도.

3. 출력 형식

- 출력은 단일 JSON 배열로 제한(설명/텍스트 금지)
- 샘플 예시를 포함해 프롬프트에 형식 고정

페르소나 JSON 구조

1. 주요 구성요소

구분	항목	설명	예시
메타 데이터	PERSONA_ID	고유 식별자	"P00001", "P00002"
정형적 파라미터	Attributes	인구통계학·행동학적 특성 집합	{"age": "30대", "gender": "여성", "occupation": "직장인", ...}
정성적 속성	Purchase_probability	제품 구매 전환 가능성을 나타내는 확률 값 (0~1)	0~1
	Base_purchase_frequency_per_month	월 평균 구매 빈도(시뮬레이션에서 단위 확률로 변환)	0~1
	Monthly_propensity_modifier	12의 배열로, 월별 계절성·프로모션·명절 효과 반영	[1.0, 1.2, 1.3, 0.9, ...]

2. 구조적 특징

- 일관성: 모든 페르소나가 동일한 스키마를 따르므로 병렬적 시뮬레이션이 가능
- 유연성: 속성 카테고리는 확장 가능, 정량 파라미터는 모델 입력으로 직접 사용 가능
- 해석가능성: "누가, 왜, 얼마나 자주" 구매하는지를 데이터 구조만으로 설명 가능

프롬프트, 페르소나 구성 결론

프롬프트 설계



- 페르소나 JSON은 단순한 텍스트 결과물이 아닌, 시뮬레이션 가능한 소비자 행동 모델로 기능
- 예측 결과는 **설명가능성(explainability)**과 **실행가능성(actionability)**을 동시에 갖춘 데이터 자산으로 활용 가능

페르소나 생성 시 사용한 프롬프트 템플릿

제품 키워드	예상 고객층 설명
덴마크 하이그릭요거트	20-40대 여성. 건강과 자기관리에 관심이 많고, 소득 수준 중상 이상인 1인 가구. 주로 온라인 채널을 통해 건강식품 구매.
동원맛참 고소참기름	20-40대 남녀: 편의성과 새로운 맛 추구, SNS(유튜브·인스타) 활용, 1인 가구 학생 및 직장인 중심. 30-50대 주부, …
동원맛참 매콤참기름	위 특징 + 매운맛 선호 성향 뚜렷
리챔 오믈레햄	30-50대 주부. 3인 이상 가구의 식사 책임, 대형마트 이용, 명절 등 가족 요리를 중요시
동원참치액 수	30-60대 주부 또는 1인 가구. 집밥 선호, 깊은 맛을 내기 위한 조미료에 투자. 대형마트와 온라인 채널 모두 이용
동원참치액 진/프리미엄	위 특징 + 요리 실력이 뛰어나고, 소득 수준이 높아 프리미엄 제품을 선호하는 미식가 성향
소화가 잘되는 우유 카페라떼	20-50대. 유당불내증이 있거나 소화 건강에 신경 쓰는 소비층. 일반 유제품 대신 락토프리 제품 선택, 출근길·점심시간에 편의점 구매 비중 높음
소화가 잘되는 우유 바닐라라떼	위 특징 + 단맛을 선호하는 젊은 층 비중 높음

- 시장 조사를 통해 제품별 예상 타겟 프로필을 설정
- 예상 고객층에 대한 설명을 포함한 동적 프롬프트를 사용하여 월별 소비 패턴을 가진 페르소나를 생성
- 생성된 페르소나들이 제품 맥락에 맞게 논리적 일관성을 유지하게 함

페르소나에게 부여된 속성

속성	값
age (연령대)	10대, 20대, 30대, 40대, 50대, 60대 이상
gender (성별)	남성, 여성
occupation (직업)	직장인, 학생, 주부, 프리랜서, 자영업자, 무직, 은퇴
income_level (소득수준)	상, 중상, 중, 중하, 하
household_size (가구 규모)	1인 가구, 2인 가구, 3인 가구, 4인 가구 이상
lifestyle (라이프스타일)	건강지향, 자기관리, 편의성추구, 가성비중시, 트렌드추구, 요리애호가, 집밥선호, 미식가
media_consumption (미디어 소비)	YouTube, Instagram, TV, 커뮤니티사이트 등
price_sensitivity (가격 민감도)	높음, 중간, 낮음
brand_loyalty (브랜드 충성도)	높음, 중간, 낮음
dietary_preferences (식습관/취향)	고단백, 저칼로리, 유당불내증케어, 소화편한음식선호, 매운맛선호 등
shopping_channel (구매 채널)	대형마트, 온라인, 편의점, 백화점

- LLM을 통해 페르소나에게 연령대, 성별, 직업 등의 속성 부여
- 제품별 30개의 페르소나 생성
- 각각의 페르소나는 속성 외에 구매 확률, 구매 빈도, 월별 소비 가중치를 가짐
- 속성들을 통해 “누가 어떤 성향을 가지고 어떤 채널을 통해 구매하는가”를 설명 가능한 형태로 변환

예측 파이프라인 개요

구분	설명
문제 정의	제품별로 “누가, 왜, 언제, 얼마나“ 살지를 정량적으로 예측하기 위해서는 정성적 소비자 맥락을 정량 모델이 읽을 수 있는 구조로 변환
핵심 아이디어	LLM이 만들어내는 소비자 묘사를 표준화된 JSON으로 강제한 결과를 에이전트 기반 시뮬레이션(ABM)에 바로 투입
결과	페르소나는 단순한 텍스트가 아니라, 모델에 바로 입력이 가능한 즉, ‘행동 가능한 데이터’

“누가, 왜, 어떻게”의 스토리를 숫자로 계산 가능한 형태로 바꾸고, 설명가능성과 실행가능성을 동시에 확보

2. 예측값 산출 시뮬레이션

판매수량 분석 핵심 요소

시뮬레이션에 사용한 기반 모델

핵심 알고리즘이 적용된 코드

코드로 구현한 시뮬레이션

핵심 알고리즘

Bass Diffusion Model

Agent Based Model

TAM, 시장점유율, 월별
가중치를 이용한 계산식

Bass Diffusion Model

- 신제품이 시장에서 어떻게 확산되는지를 설명하는 이론적 모델
- 소비자는 ‘혁신가’와 ‘모방가’ 두 그룹으로 나누어 시장 잠재 구매자의 누적 채택률을 예측
 - 혁신가(Innovator): 외부 정보(광고, 홍보 등)에 영향을 받아 제품을 채택
 - 모방가(Imitator): 이미 제품을 구매한 사람들(입소문, 사회적 전염 등)의 영향을 받아 채택
- 위 두 가지 요소를 사용하여 특정 시점의 누적 채택 확률을 계산
- 매월 새롭게 제품을 채택하는 고객 수를 예측 및 시뮬레이션 가능

참고 논문: Bass, F. M. (1969). A new product growth for model consumer durables. Management Science, 15(5), 215-227. [31]

Agent Based Model

- 복잡한 시장을 개별 에이전트들의 행동 규칙과 상호작용으로 이해하려는 시뮬레이션 기법
- 본 과제에서는 페르소나가 에이전트 역할을 하게 됨
- 각 페르소나는 비인지(Unware), 인지(Aware), 활성 고객(Active), 이탈 고객(Churned)과 같은 상태를 가짐
- 그리고 매월 특정 확률에 따라 상태를 바꾸고, 전체 시장의 판매량은 이런 개별 에이전트들의 행동이 모여 나타나는 결과
물로 예측해볼 수 있음
- 복잡한 시장 행동의 움직임을 최대한 비슷하게 구현하는것에 초점을 둠

참고 논문: Epstein, J. M., & Axtell, R. L. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. The MIT Press. [32]

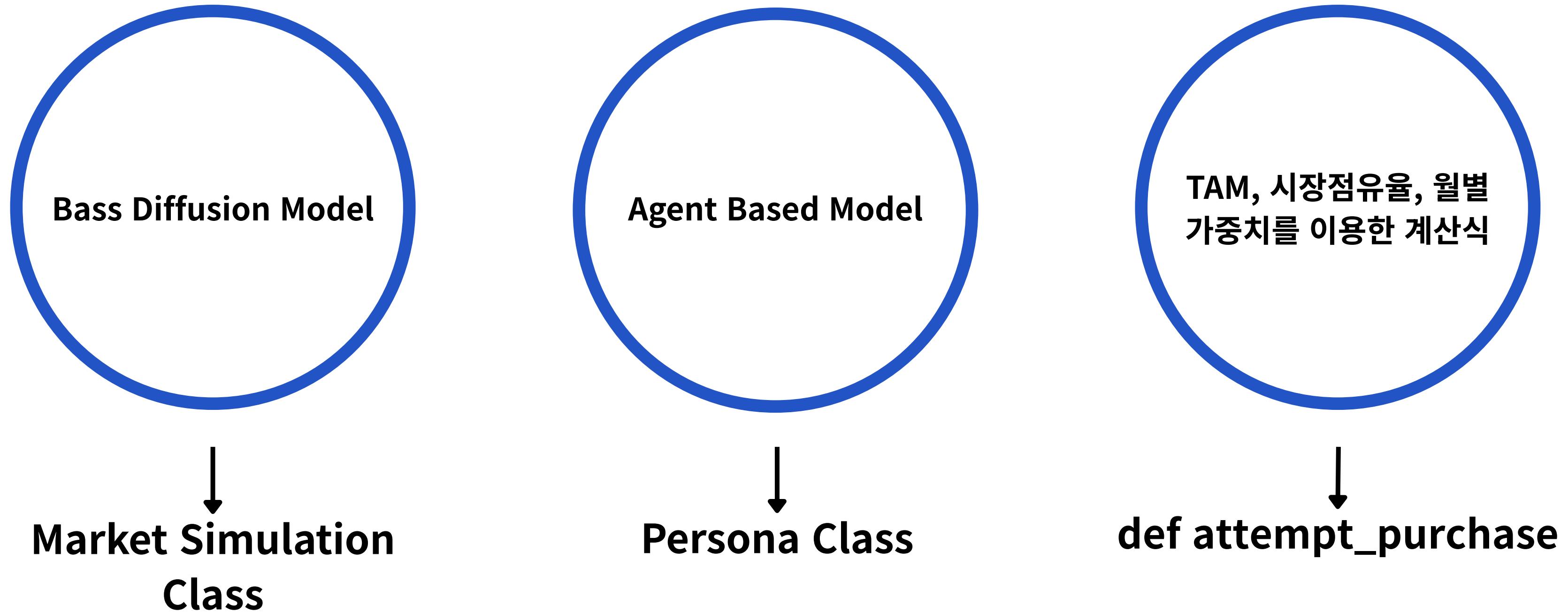
TAM, 시장점유율, 월별 가중치를 이용한 계산식

$$P_{MonthlyPurchase} = (1 - (1 - P_{Daily})^{30}) \times Modifier_{Month}$$

- **P_MonthlyPurchase:** 한 달 동안 최소 한 번 이상 구매할 최종 확률
- **P_Daily:** 하루에 구매할 기본 확률
- **Modifier_Month:** 월별 계수

페르소나의 `base_purchase_frequency_per_month` 의 값을 30으로 나누어 하루에 구매할 확률을 구하고, 이 값으로 하루에 구매를 안 할 확률을 30번 제곱해 한 달 동안 구매를 한 번 이상 할 확률을 구함
월별 계수로 곱해주어 나온 숫자와 랜덤한 숫자를 비교해 위의 식이 더 큰 수이면 구매 성공을 의미

핵심 알고리즘이 적용된 코드



코드로 구현한 시뮬레이션

Persona Class

- 이 클래스는 개별 소비자 에이전트를 시뮬레이션하기 위한 구조
- Gemini가 생성한 페르소나 JSON파일의 내용을 입력으로 받아 변수로 저장
- Agent Based Model에 기반하여 비인지, 인지 등의 상태 변수를 가지고 있음
- Active 상태의 에이전트는 확률적으로 Churned 상태로 변환
- Unware에서 Active로 전환될 확률은 다음 Bass Diffusion Model 공식에 따름
 - $P(\text{aware}) = p_{\text{innovation}} + p_{\text{imitation}} \times \text{adoption_rate}$
- 또한, Active 상태의 에이전트는 계산된 확률값을 사용하여 구매 여부를 결정
- 다음 확률 계산을 기반으로 한 달 내 구매 여부 결정
 - $P(\text{monthly purchase}) = (1 - (1 - \text{base_purchase_rate})^{30}) \times \text{month_modifier}$
 - 난수에 기반하여 설정한 Threshold에 따라 구매 여부에 변칙성을 부여함(실제)

코드로 구현한 시뮬레이션

Market Simulation Class

- 시장의 전체적인 시뮬레이션을 담당하는 클래스
- 시장 규모, 시장 점유율, 월별 계절성 보정 계수 리스트, 초기 채택률 등을 생성 인자로 받음
- 페르소나 JSON과 클래스를 사용해서 시뮬레이션을 위한 페르소나 생성
- 초기 채택률에 따라 일부 에이전트를 Active 상태로 설정
- run_simulation(months=12) 함수를 통해 월 단위 시장의 구매 시뮬레이션을 수행
 - 채택률 계산 : $\text{adoption_rate} = \text{현재 활성 에이전트 수} / \text{전체 시장 잠재 수요 수}$
 - 각 에이전트 순회
 - 상태 전이(update_state) 수행
 - 구매 시도(attempt_purchase)
 - Active 상태 수 추적
 - 구매율 추정: 샘플 에이전트 기준 구매율 x 전체 시장 잠재 수요 추정 → 전체 매출 수치 도출

코드로 구현한 시뮬레이션

핵심 알고리즘 요약

1. Bass Diffusion Model 기반으로 상태 전이

- Unware → Active 전이 확률: $\text{prob_aware} = p_{\text{innovation}} + p_{\text{imitation}} \times \text{adoption_rate}$
- `np.random.rand()`로 무작위값을 뽑아 위 확률보다 작으면 Active 상태로 전환
- Active → Churned 전이는 $p_{\text{churn}} = 0.05$ 의 고정 확률로 발생

2. Active 상태에서 개별 구매 확률을 시뮬레이션

- 하루당 구매 확률을 바탕으로, 한 달 내 1회 이상 구매할 확률 계산
 - $\text{monthly_purchase_prob} = (1 - (1 - \text{base_rate})^{30}) \times \text{month_modifier}$
- Active 상태일 때에만 구매 시도, 랜덤보다 크면 실제 구매로 간주

코드로 구현한 시뮬레이션

핵심 알고리즘 요약

3. 샘플 기반 전체 시장 매출을 추정

- 에이전트 샘플 내 구매 성공률로 전체 시장 매출 예측
- $\text{extrapolated_sales} = (\text{current_month_sales} / \text{num_agents}) \times \text{potential_market_size}$
- 매달 이 수치를 저장하여 월별 시뮬레이션 결과 생성

4. 채택률을 매달 갱신하는 반복 루프

- 현재 활성화된 샘플의 비율로 전체 시장 기준 채택자 수 추정
 - $\text{adopters} = (\text{active_agents} / \text{num_agents}) \times \text{potential_market_size}$
- 이 수치는 다음 달의 확산 확률(adoption_rate) 계산에 사용됨

3. 코드 실행 방법

함수 및 파라미터 설명

메인 시뮬레이션 루프

코드 실행 방법

함수 및 파라미터 설명

```

class PersonaAgent:
    """
    개별 소비자 에이전트를 시뮬레이션하기 위한 클래스
    Agent Based Model에 기반하여 비인지, 인지 등의 상태 변수를 가지고 있습니다.
    Active 상태의 에이전트는 확률적으로 Churned 상태로 변환
    Unaware에서 Active로 전환될 확률은 다음 Bass Diffusion Model 공식에 따릅니다.
    또한, Active 상태의 에이전트는 계산된 확률값을 사용하여 구매 여부를 결정합니다.
    """

class MarketSimulation:
    """
    시장의 전체적인 시뮬레이션을 담당하는 클래스
    시장 규모, 시장 점유율, 월별 계절성 보정 계수 리스트, 초기 채택률 등을 생성 인자로 받습니다.
    페르소나 JSON과 클래스를 사용해서 시뮬레이션을 위한 페르소나 생성합니다.
    초기 채택률에 따라 일부 에이전트를 Active 상태로 설정합니다.
    run_simulation(months=12) 함수를 통해 월 단위 시장의 구매 시뮬레이션을 수행합니다.
    """

def apply_holiday_gift_set_boost(product_name, monthly_sales, logger):
    """
    특정 제품에 대해 명절(2월, 9월) 회사차원 선물 세트 판매량을 추가합니다.
    - product_name (str): 제품명
    - monthly_sales (list): 시뮬레이션으로 예측된 12개월 판매량
    - logger (Logger): 로깅 객체
    Returns:
    """

```

함수 및 파라미터 설명

페르소나 생성 프롬프트 함수

```
def create_product_specific_prompt(product_name, num_personas=30):  
    """
```

상품명과 생성할 페르소나 수를 입력으로 받아 페르소나 생성에 사용할 프롬프트를 반환하는 함수

```
"""
```

최종 판매량 숫자 정리 함수

```
def format_sales_numbers(product_name, monthly_sales, logger):  
    """
```

제품군별 규칙에 따라 최종 판매량 숫자를 반올림하여 정리합니다.

- product_name (str): 제품명
- monthly_sales (list): 모든 계산이 끝난 12개월 판매량
- logger (Logger): 로깅 객체

Returns:

- list: 반올림 규칙이 적용된 12개월 판매량

```
"""
```

함수 및 파라미터 설명

시뮬레이션 파라미터 정의

제품 목록

```
ESTABLISHED_PRODUCTS = [
    '동원맛참 고소참기름 135g', '동원맛참 고소참기름 90g', '동원맛참 매콤참기름 135g', '동원맛참 매콤참기름 90g',
    '동원참치액 순 500g', '동원참치액 순 900g', '동원참치액 진 500g', '동원참치액 진 900g',
    '프리미엄 동원참치액 500g', '프리미엄 동원참치액 900g'
]
```

제품 출시월

```
NEW_PRODUCT_LAUNCH_DATES = {
    '덴마크 하이그릭요거트 400g': (2025, 2),
    '리챔 오믈레햄 200g': (2025, 5),
    '리챔 오믈레햄 340g': (2025, 5),
    '소화가 잘되는 우유로 만든 바닐라라떼 250mL': (2025, 2),
    '소화가 잘되는 우유로 만든 카페라떼 250mL': (2025, 2)
}
```

제품 출시월

```
HOLIDAY_MODIFIERS = {
    'seollal_chuseok': [2.4, 2.2, 0.9, 1.0, 1.1, 1.0, 1.0, 2.4, 2.2, 0.8, 0.9, 0.8],
}
```

```
DEFAULT_MODIFIERS = [1.0] * 12
```

함수 및 파라미터 설명

시뮬레이션 파라미터 정의

제품별 시뮬레이션 가중치

```
SIMULATION_PARAMS = {
    '덴마크 하이그릭요거트 400g': {'tam': 8000000, 'market_share': 0.04, 'modifiers': [1.2, 1.1, 1.6, 1.8, 1.8, 1.9, 1.2, 1.2, 0.9, 0.8, 0.9, 0.9]},
    '동원맛참 고소참기름 135g': {'tam': 20000000, 'market_share': 0.043, 'modifiers': [0.7, 0.7, 0.9, 1.0, 1.0, 1.2, 1.4, 1.1, 0.6, 0.7, 0.6, 0.7]},
    '동원맛참 고소참기름 90g': {'tam': 20000000, 'market_share': 0.055, 'modifiers': [0.7, 0.7, 0.9, 1.0, 1.0, 1.2, 1.4, 1.1, 0.6, 0.7, 0.6, 0.7]},
    '동원맛참 매콤참기름 135g': {'tam': 20000000, 'market_share': 0.033, 'modifiers': [0.7, 0.7, 0.9, 1.0, 1.0, 1.2, 1.4, 1.1, 0.6, 0.7, 0.6, 0.7]},
    '동원맛참 매콤참기름 90g': {'tam': 20000000, 'market_share': 0.04, 'modifiers': [0.7, 0.7, 0.9, 1.0, 1.0, 1.2, 1.4, 1.1, 0.6, 0.7, 0.6, 0.7]},
    '리챔 오믈레햄 200g': {'tam': 5500000, 'market_share': 0.065, 'modifiers': [0.9, 0.8, 1.2, 1.2, 1.2, 1.1, 1.2, 1.3, 0.9, 1.1, 1.0, 1.1]},
    '리챔 오믈레햄 340g': {'tam': 5500000, 'market_share': 0.04, 'modifiers': [0.9, 0.8, 1.2, 1.2, 1.2, 1.1, 1.2, 1.3, 0.9, 1.1, 1.0, 1.1]},
    '동원참치액 순 500g': {'tam': 2500000, 'market_share': 0.042, 'modifiers': [2.4, 2.2, 0.9, 1.0, 1.1, 1.0, 1.0, 2.4, 2.2, 0.8, 0.9, 0.8]},
    '동원참치액 순 900g': {'tam': 2500000, 'market_share': 0.019, 'modifiers': [2.4, 2.2, 0.9, 1.0, 1.1, 1.0, 1.0, 2.4, 2.2, 0.8, 0.9, 0.8]},
    '동원참치액 진 500g': {'tam': 2500000, 'market_share': 0.06, 'modifiers': [2.4, 2.2, 0.9, 1.0, 1.1, 1.0, 1.0, 2.4, 2.2, 0.8, 0.9, 0.8]},
    '동원참치액 진 900g': {'tam': 2500000, 'market_share': 0.025, 'modifiers': [2.4, 2.2, 0.9, 1.0, 1.1, 1.0, 1.0, 2.4, 2.2, 0.8, 0.9, 0.8]},
    '프리미엄 동원참치액 500g': {'tam': 2500000, 'market_share': 0.014, 'modifiers': [1.3, 1.0, 1.0, 1.0, 1.1, 0.9, 0.8, 0.9, 1.4, 0.9, 1.2, 1.2]},
    '프리미엄 동원참치액 900g': {'tam': 2500000, 'market_share': 0.003, 'modifiers': [1.1, 1.0, 1.0, 1.0, 1.1, 0.9, 0.9, 0.9, 1.1, 0.9, 1.2, 1.1]},
    '소화가 잘되는 우유로 만든 바닐라라떼 250mL': {'tam': 16000000, 'market_share': 0.11, 'modifiers': [0.9, 0.9, 1.0, 1.1, 1.1, 1.1, 1.2, 1.2, 1.1, 1.1, 1.1, 1.0]},
    '소화가 잘되는 우유로 만든 카페라떼 250mL': {'tam': 16000000, 'market_share': 0.11, 'modifiers': [0.9, 0.9, 1.0, 1.1, 1.1, 1.1, 1.2, 1.2, 1.1, 1.1, 1.1, 1.0]},
    'default': {'tam': 10000000, 'market_share': 0.10, 'modifiers': DEFAULT_MODIFIERS}
}
```

메인 시뮬레이션 루프

1. 샘플 파일 불러오기 및 최종 제출 파일 경로 설정

```
try:
    submission_df = pd.read_csv(os.path.join(PATH, 'sample_submission.csv'))
    logger.info("✓ 4. 제출용 데이터프레임 로드를 완료했습니다.")
except FileNotFoundError:
    logger.error(f"! 'sample_submission.csv' 파일을 찾을 수 없습니다.")
    exit()
```

```
PERSONAS_PER_BATCH = 30
NUM_BATCHES_PER_PRODUCT = 10
MAX_ATTEMPTS_PER_BATCH = 3
submission_filename = os.path.join(PATH, f'submission final {timestamp}.csv')
```

2. 시뮬레이션 루프 시작

```
for index, row in submission_df.iterrows():
    product_name = row['product_name']
    logger.info(f"\n===== [ {product_name} ] 판매량 예측 시작 =====")

    params = SIMULATION_PARAMS.get(product_name, SIMULATION_PARAMS['default'])
```

메인 시뮬레이션 루프

3. 생성된 페르소나 JSON 파일 확인 후 불러오기

```
if not USE_API_TO_GENERATE_PERSONAS and os.path.exists(persona_cache_file):
    try:
        with open(persona_cache_file, 'r', encoding='utf-8') as f:
            product_personas = json.load(f)
        logger.info(f"✓ [캐시 사용] '{persona_cache_file}'에서 페르소나 {len(product_personas)}명을 성공적으로 불러왔습니다.")
    except Exception as e:
        logger.warning(f"! 캐시 파일 '{persona_cache_file}' 로드 중 오류 발생: {e}. API를 통해 재생성을 시도합니다.")
    product_personas = []
```

4. 페르소나가 없는 경우 Gemini를 통해 생성

```
if not product_personas:
    if USE_API_TO_GENERATE_PERSONAS:
        logger.info("📝 [API 모드] Gemini API를 통해 페르소나 생성을 시작합니다.")
        for i in range(NUM_BATCHES_PER_PRODUCT):
            for attempt in range(MAX_RETRIES_PER_BATCH):
                try:
                    prompt = create_product_specific_prompt(product_name, PERSONAS_PER_BATCH)
                    logger.info(f"⏳ 배치 {i+1}/{NUM_BATCHES_PER_PRODUCT} API 호출 중... (시도 {attempt+1})")
                    response = model.generate_content(prompt)

                    json_text = extract_json_from_response(response.text)
                    if not json_text:
                        raise ValueError("응답에서 JSON을 찾을 수 없습니다.")
```

메인 시뮬레이션 루프

5. Market Simulation Class를 사용하여 시뮬레이션 진행

```
if product_name in ESTABLISHED_PRODUCTS:  
    initial_rate = np.random.uniform(0.45, 0.5)  
else:  
    initial_rate = np.random.uniform(0.03, 0.05)  
  
logger.info(f"--- [ {product_name} ] ABM & Bass Model 시뮬레이션 시작 ---")  
market_sim = MarketSimulation(  
    personas=product_personas,  
    tam=params['tam'],  
    market_share=params['market_share'],  
    modifiers=params['modifiers'],  
    initial_adoption_rate=initial_rate  
)  
  
monthly_sales = market_sim.run_simulation(months=12)  
  
monthly_sales = apply_holiday_gift_set_boost(product_name, monthly_sales, logger)
```

코드 실행 방법

함께 제출한 페르소나 파일을 사용할 경우 생략 가능

Gemini API 사용을 위한 API Key 설정

USE_API_TO_GENERATE_PERSONAS 변수로 기존에 생성한 페르소나 재사용 설정

```
if USE_API_TO_GENERATE_PERSONAS:  
    try:  
        genai.configure(api_key="Use your API Key")  
        model = genai.GenerativeModel('models/gemini-2.5-flash')  
        logger.info("✓ Gemini API 키가 설정되었습니다. [API 모드]")  
    except Exception as e:  
        logger.error(f"! API 키 설정 중 오류가 발생했습니다: {e}")  
        USE_API_TO_GENERATE_PERSONAS = False  
        logger.warning("! API 사용이 불가능하여 [캐시 사용 모드]로 강제 전환합니다.")
```

코드 실행 방법

폴더 구조 설명 및 실행 요청

📦 프로젝트 루트

```
├── data
│   └── sample_submission.csv
└── personas
    ├── 덴마크 하이그릭요거트 400g_personas.json
    ├── 동원맛참 고소참기름 90g_personas.json
    ├── 동원맛참 고소참기름 135g_personas.json
    ├── 동원맛참 매콤참기름 90g_personas.json
    ├── 동원맛참 매콤참기름 135g_personas.json
    ├── 동원참치액 순 500g_personas.json
    ├── 동원참치액 순 900g_personas.json
    ├── 동원참치액 진 500g_personas.json
    ├── 동원참치액 진 900g_personas.json
    └── ...
```

main.py

개발 환경

Python==3.10
google-generativeai==0.8.5
pandas==2.3.2
numpy==2.2.6

API Key 설정 또는 persona 폴더 확인 후
모든 실행 코드가 담긴 main.py를 실행

4. 제품별 시장 분석

판매수량 분석 핵심 요소

덴마크 하이 그릭요거트 분석

동원맛참 분석

리챔 오믈레햄 분석

동원참치액 분석

소화가 잘되는 라떼 분석

판매수량 분석 핵심 요소

시장 규모

연간 총매출로 시장 크기 계량화

→ 예측 대상 시장의 현실적 범위 파악

핵심 지표 : 연간 카테고리 매출, 채널별 매출, CAGR

시장 점유율

소비자가 선택할 수 있는 제품 중

동원 제품의 상대적 판매량 비율 판단

소비자 규모

타겟 고객의 인구통계, 심리, 행동 특성을 통해 잠재 구매자 전체 (TAM) 산정

핵심 지표 : 인구수 × 침투율 × (구매자 1인당 구매빈도/구매량)

월별 소비 패턴

연중 수요 변동성을 분석해 월별 가중치로 계량화. (성/비수기 반영)

핵심 지표 : 월별 판매지수, 계절·프로모션 효과.

하이 그릭 요거트 시장 규모

그릭 요거트 시장의 경우 닐슨리서치 조사에 따르면 2024년 기준 약 **900억원** 규모 [2]

웰니스와 다이어트 트렌드 확산을 계기로 국내 그릭요거트 시장도 수혜를 입은 것으로 분석
팬데믹 시기에는 홈디저트 수요와 프로바이오틱 등 건강 기능 성에 대한 관심이 높아지면서 요거트 시장 전체가 확대됨

그릭 요거트 시장이 앞으로도 **꾸준한 성장세**를 보일 것으로 판단

- 고단백, 제로슈거, 락토프리와 같은 프리미엄 기능성 제품이 신제품 출시의 주류를 형성하며 시장 재편을 이끌고 있음
- 이러한 흐름은 소비자들의 헬시플레저 트렌드, 락토프리 인구 증가, 프리미엄 식품 선호 확대와 맞물려 향후에도 국내 그릭요거트 시장의 꾸준한 성장세를 예상 [3]



하이 그릭 요거트 소비자 수

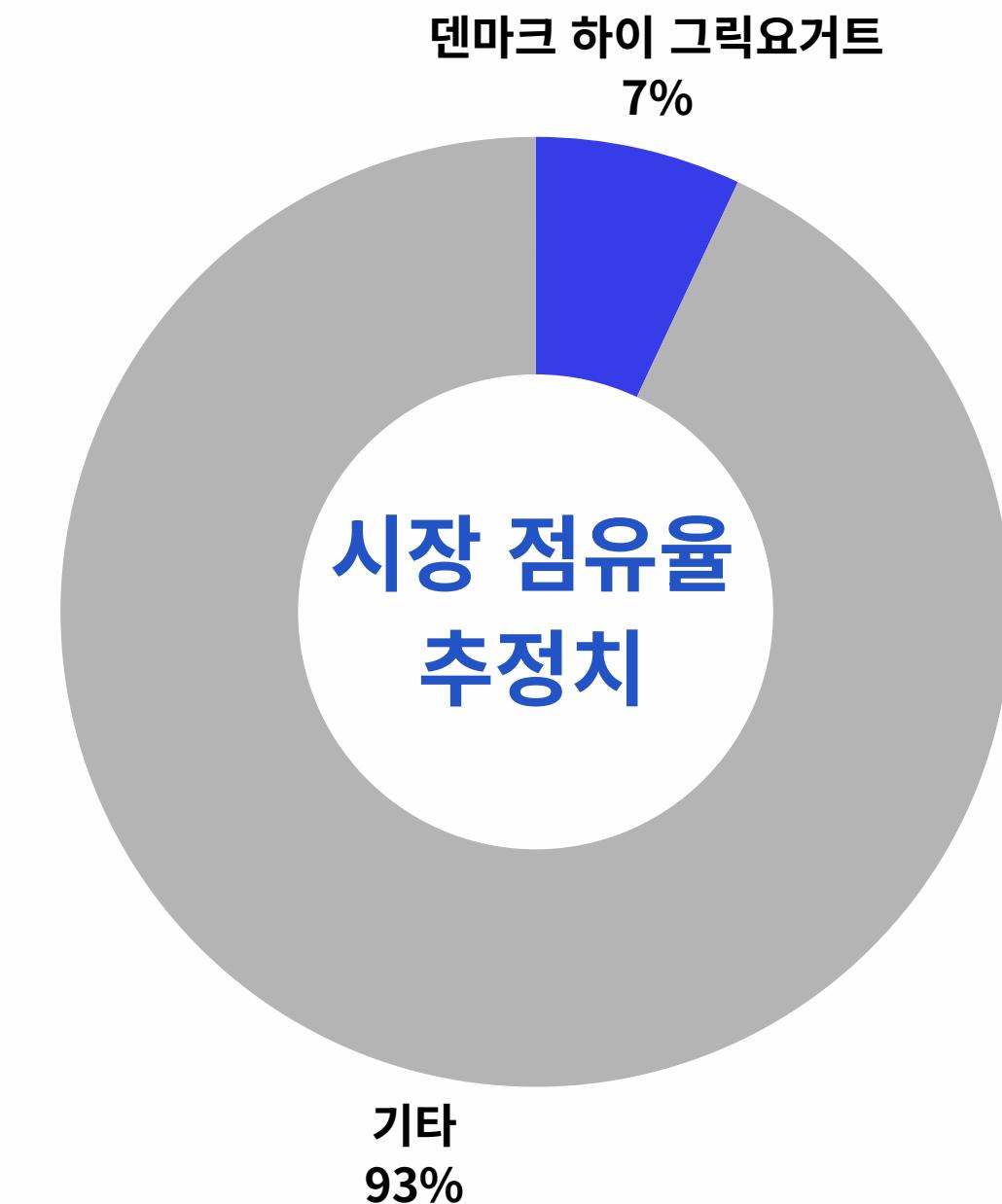
800 만명

20~40대 소비자
2000만명 중 35% + α

타겟 고객층은 20~40대 인구, 그 외엔 헬스 피트니스 소비층, 락토프리 소비층이 영향을 줄것으로 판단

- 20~40대 인구는 2000만명이며, 여성의 경우 다이어트 및 체형 관리 니즈가 크고, 저칼로리, 고단백인 그릭요거트를 간식 또는 한끼 대체식으로 소비 [4]
 - 결혼한 경우 어린이 아침 대체식으로 활용을 위해 더 안정적인 수요를 보일 것으로 예상
 - 꾸준하게 그릭요거트를 소비할 인구는 유산소 신체활동 실천율과 유사한 관계를 보일 것으로 판단, 실천율은 약 35% 추정[5]
 - $2000 * 0.35 = 700\text{만명}$
- 헬스 피트니스 소비층 중에 식단에 신경을 많이 쓰는 소비층은 단백질 보충 목적이 뚜렷하며, 프로틴바, 보충제와 함께 운동 후 간편 섭취하는 경우가 많음. [6]
- 유당불내증 인구 중 기존 발효유를 대체할 제품으로 수요 증가하는 것도 고려. 필수층이지만 성장 속도는 상대적으로 완만 [7]

하이 그릭 요거트 시장 점유율



동원F&B의 프리미엄 유가공 브랜드 ‘덴마크 하이’는 2023년 2월 론칭 후 **1년 만에 누적 매출 200억 원을 돌파하며, 프리미엄·락토프리·고단백이라는 강점을 기반으로 건강 지향 소비층의 빠른 지지를 얻음**

- 단순한 매출 성과를 넘어, 단기간 내에 시장 내 빠른 침투력과 소비자 기반 형성을 통해 브랜드 인지도를 확보함 [8]
- 발효유 시장에서 그릭요거트 시장은 일부지만 빠르게 확보인 인지도와 유당 zero라는 키워드로 인해 2025년 2월 출시부터 상당한 관심을 받았을 거라 판단
- 덴마크 하이의 제품들 사이에서 33%의 매출을 올렸다고 가정할 때 연매출 66억은 그릭요거트 시장 900억의 7%
- 그릭요거트 시장에서의 점유율은 **4~10%** 사이에 있을 것으로 추측

하이 그릭 요거트 월별 소비 패턴

- **1~2월 (저조기)**
- 건강관리 니즈는 있으나 명절 전후 소비 위축으로 실제 수요는 낮음.
- **3~6월 (성수기) [9]**
- 봄맞이 체형 관리·헬시플레저 트렌드와 여름 대비 단백질 보충 수요가 결합해 소비가 가장 활발.
- “고단백·락토프리·프리미엄 간식” 이미지가 성장을 견인.
- **7~8월 (보통기)**
- 휴가철 이후 다이어트 긴장감은 완화되지만, 헬스족·건강 간식 선호층이 꾸준히 소비를 유지.
- 일부 고객은 더운 날씨 속 시원한 간편 간식으로 선택.
- **9~12월 (비수기)**
- 연말 고칼로리 식습관 영향으로 수요 감소. 다만 프리미엄 건강 간식·락토프리 대체식품을 찾는 틈새 소비층은 유지.

덴마크 그릭요거트

1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
1.2	1.1	1.6	1.8	1.8	1.9	1.2	1.2	0.9	0.8	0.9	0.9

동원맛참 시장 규모

닐슨아이큐코리아(NIQ)에 따르면 소매점 매출 기준 2023년 3366억, 2024년 **3458억** 규모로 전년동기 대비 2.73% 증가

소비 트렌드가 온라인으로 이동하면서 온라인 매출은 2020~2023년까지 연평균 20% 이상 성장, 오프라인을 포함한 전체

매출도 꾸준히 성장하고 있음 [10]

1인 가구를 중심으로 간편하고 저렴한 식사를 찾는 소비자들을 중심으로 참치캔은 꾸준히 팔려 나가고 있고, 이런 추세에 맞춰 별도 조리 없이 바로 식사용으로 먹을 수 있게 개공한 참치캔 매출 또한 함께 성장세를 보일 것으로 판단됨



동원맛참 소비자 수

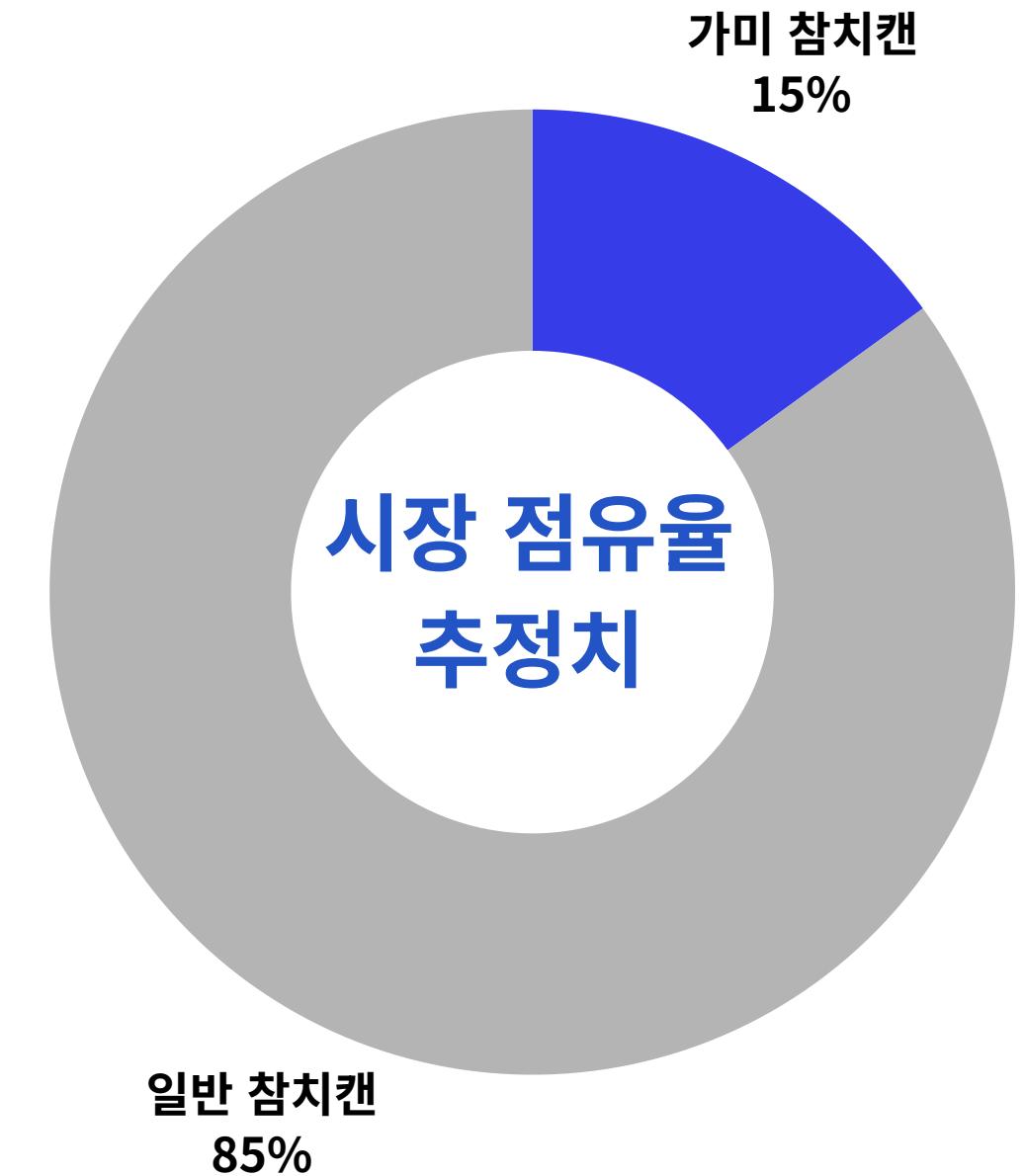
2000만명 추정

20-40대 인구 2000만명 $\pm \alpha$

참치캔의 주 타겟층은 **20-40대 남녀**, 동원 참치캔에 대한 인식은 대중적이므로 타겟층 전부를 소비자 수로 판단

- 전통적인 참치캔 소비자는 전부 가미 참치캔의 잠재적 소비자로 산정할 수 있음
- 편의성과 새로운 맛 추구하는 1인 가구 학생 및 직장인 중심으로 소비가 이뤄질 것으로 예상
- 일상적인 상황에서 반찬거리로 적당한 조미 참치캔은 접근성도 뛰어날 것으로 판단
- 2024년 건강보험 기준 20-40대 인구는 **2000만명** [11]

동원맛참 시장 점유율



- 동원의 가미 참치캔은 전체 참치캔 시장에서 별도 집계는 없지만, 전체 동원참치 점유율(**약 80%**) 중 일부를 차지
- 업계 추정치에 따르면, 가미 참치캔이 동원 내에서 차지하는 비중을 고려했을 때 전체 시장에서 **약 15~20%** 수준의 점유율을 기록할 것으로 예상
- 주요 경쟁사는 사조해표, 오뚜기 등이 있으나 가미 참치만 놓고 보면 직접적인 대체제는 제한적

동원맛참 월별 소비 패턴

동원맛참 고소참기름/매콤참기름



- 20-30대 남/녀 소비 패턴을 기반으로 판단했을 때 6-8월(여름)엔 캠핑/야외활동으로 간편식 수요가 상승할 것으로 예상
- 1-2월, 9-10월은 명절 선물용으로 인기가 많은 조리용 일반 참치캔의 수요가 증가하여 오히려 가미 참치캔의 판매량은 감소할 것이라 판단 [12]
- 11-12월(겨울)에는 주로 집에서 따뜻한 음식을 조리해서 먹는 경우가 많아져 간편식보다 식재료로서의 참치캔의 소비가 늘어나 가미 참치캔의 소비가 감소하는 경향이 있음

오믈레햄 시장 규모

캔햄은 유로모니터에 따르면 2022년 **6675억원** 정도의 시장 규모 [13]

나트륨 및 지방 함량을 낮춘 제품이나 식물성 캔햄 시장이 성장세
국산 돼지고기 함량에 대한 이슈로 캔햄의 원재료 및 성분 차이에 대한 관심이 조명됨

캔햄 시장이 2024년, 2025년 동안에 **현상유지를** 할 것이라 예상

- 캔햄과 같이 생활 식습관에 깊숙히 침투한 제품의 경우 연간 소비량이 **경제성장률과 비슷할** 것이라 예상
 - 최근 몇 년간 경제성장률이 부진
- 대회기간 동안의 시장규모는 보수적으로 **6700억원** 정도의 예상



오믈렛 소비자 수

550
만명

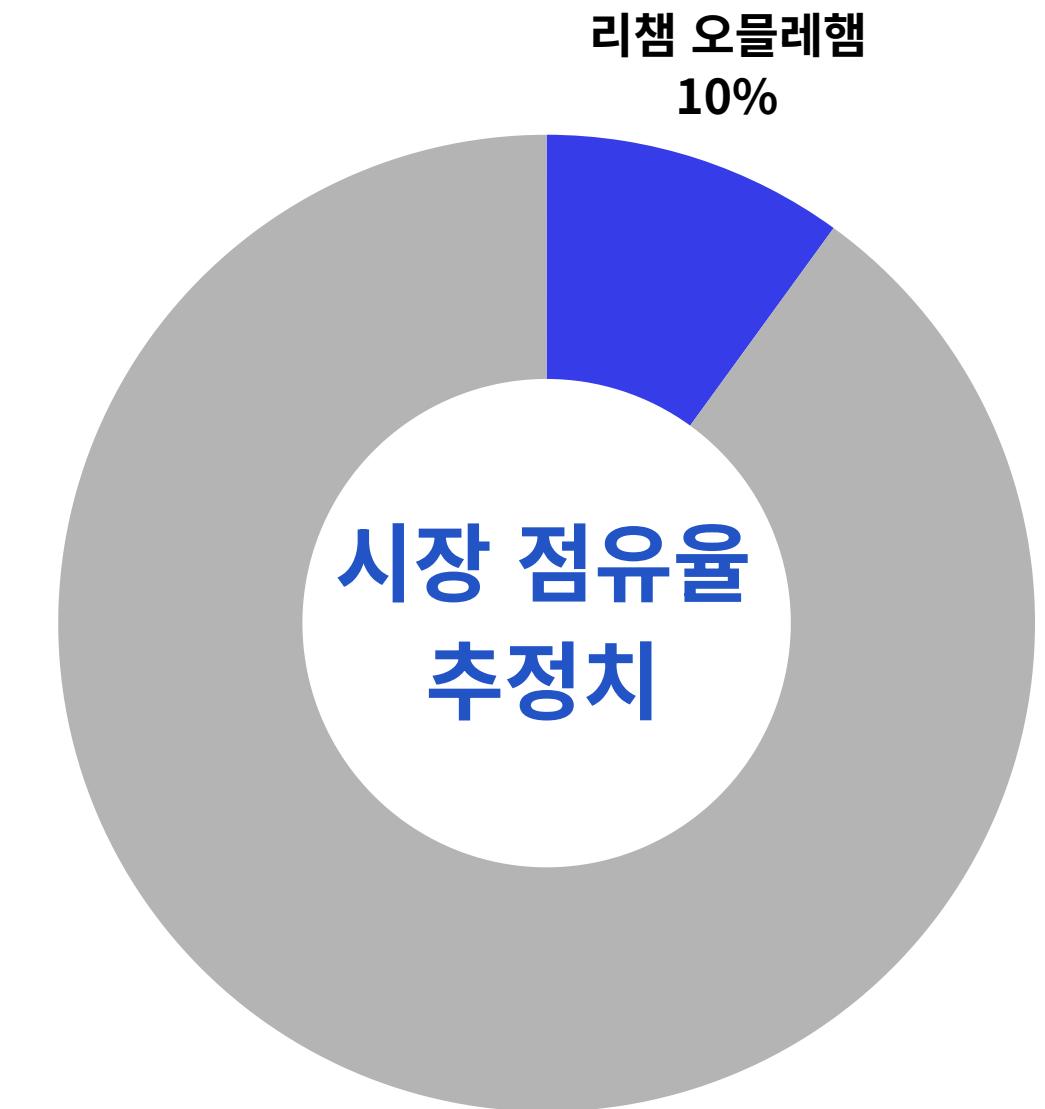
자녀와 동거하는 가구
523 만명 + α

일상적인 식사에 사용할 목적이 있는 가구에서 주로 소비할 것으로 추측 [14]

- 리챔은 꾸준히 저나트륨에 집중하는 방향성을 가짐 → 나트륨을 신경쓰는 소비자
- 영유아를 제외한 **자녀를 양육하는 가구가 소비의 주축**
 - 2024년 18세 미만 자녀 양육 가구(451만) - 영유아 양육가정(128만) [15]
 - 20~49세의 인구중 약 30퍼센트정도가 부모와 동거(600만) [16][17]
 - 20~49세가 속한 가구원의 수는 대한민국 평균 가구원수 2.2 보다 높을 것으로 추측 [18] → 3.0

$$323 + 600 / 3 + \text{alpha} = 550\text{만명}$$

오믈레햄 시장 점유율



리챔은 캔햄 시장 점유율의 20% 정도를 차지 [19]

- 그동안 쌓아온 브랜드의 인지도와 충성고객을 충분히 확보했다고 판단
- 리챔 오믈레햄만의 **특수성**이 있다 판단
- 굽기만 해도 케찹 + 계란의 맛이 느껴짐

550만의 예상 소비자 중 리챔의 충성고객과 특수성을 높게 산 소비자는 대략 10% 정도일거라 추측

오믈레햄 월별 소비 패턴

리챔 오믈레햄 500g & 900g

1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
0.9	0.8	1.2	1.2	1.2	1.1	1.2	1.3	0.9	1.1	1.0	1.1

- 2024년 7-8월, 12-1월 (여름/겨울방학): 특히 학생 자녀들이 집에 머무는 시간이 길어지면서, **반찬 및 간식용**으로 캔 햄 소비가 증가
- 2024년 9월 (명절 영향): 선물세트의 메인인 일반 리챔/스팸 소비가 늘어나, '오믈레햄'과 같은 특화 제품의 수요는 상대적으로 감소
- 그 외의 기간 (학기 중): 신학기 도시락 반찬 수요와 기말고사 기간 간편식 수요로 소폭 상승

동원참치액 시장 규모

21일 시장조사기관 닐슨아이큐코리아(NIQ)에 따르면 소매 점 매출 기준 2023년 국내 참치액 시장은 약 **527억원 규모** [20]

코로나19 팬데믹을 전후해 내식 수요가 증가하면서 조미료 시장에 성장함에 따라 참치액 시장도 수혜를 입은 것으로 분석

참치액 시장이 2024년, 2025년 동안 **소폭 상승할** 것으로 예측

- 경기침체로 소비가 주춤, 외식 물가 상승으로 점점 집밥 인구 가 늘어나며 조미료 시장은 성장하고 있음 [21]
- 대중들의 인지도가 올라가면서 시장도 함께 성장하였고, 국이 나 김치에 사용하는 등 새로운 활용 시도 또한 이어지고 있음
- 그럼에도 불구하고 기존 액젓 시장은 여전히 견고함
- 보수적으로 2024.07 ~ 2025.06까지는 **600억원대**로 추정



동원참치액 소비자 수

250 만명

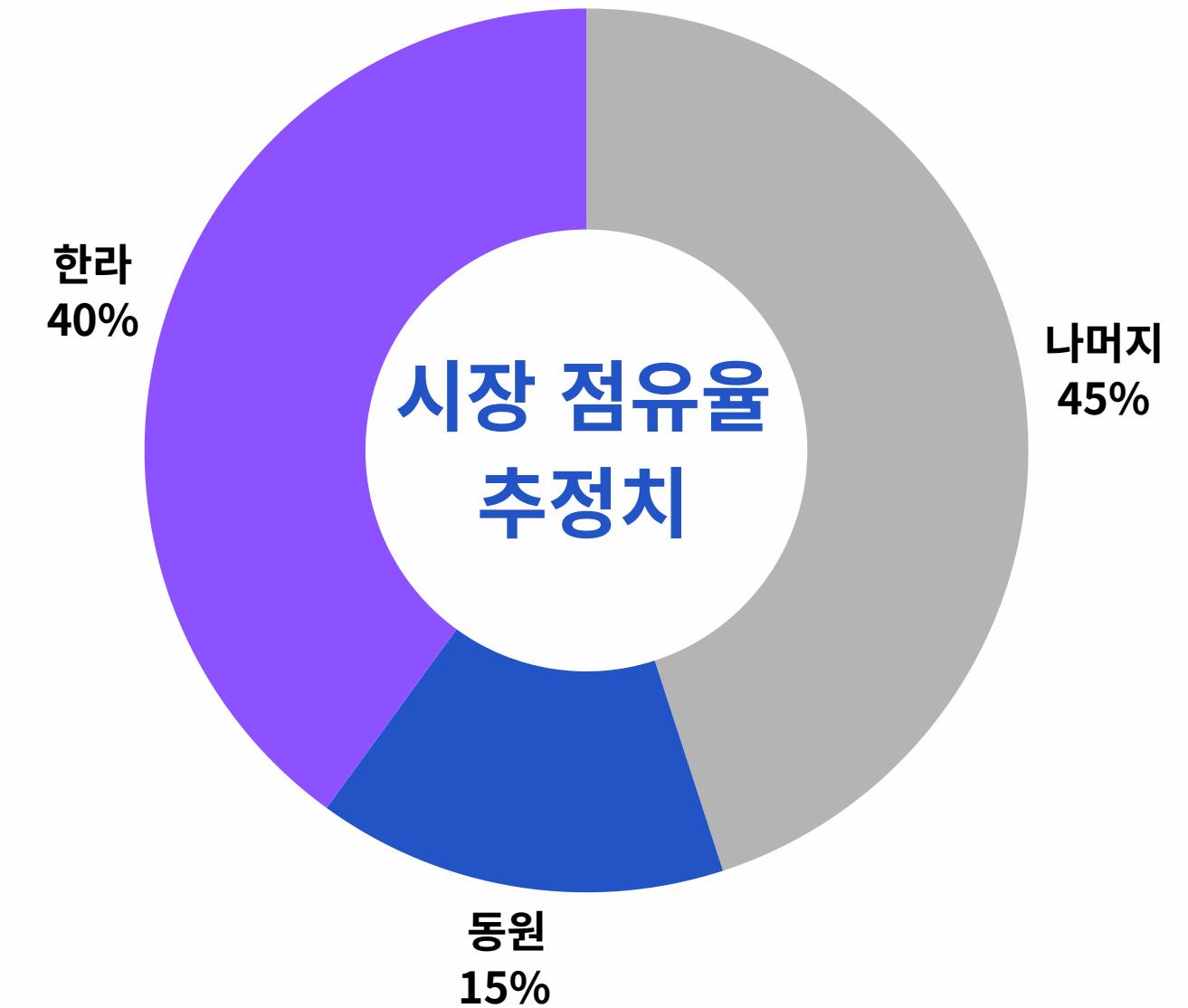
주부 700만명 중 30~40%

참치액 주 타겟층은 30-60대 주부(약 700만) 혹은 요리 애호가로 구분 [22]

- 주부 소비자들 중에서 일상적으로 집에서 요리를 하고, 특히 연령대가 높은 타겟이 주로 소비할 것이라고 예상
 - 보수적으로 주부 인구 700만의 30~40% 추정
- 그 외 요리 애호가중에서는 '국/탕/찌개' 등 한식 조리에 관심이 많은 경우
 - 요리 애호가의 숫자는 상당히 적을 것으로 예상

도합 250만 명으로 추정

동원참치액 시장 점유율



동원이나 사조에서도 적극적으로 참치액 시장에 뛰어들었으며, 크지 않은 점유율이지만 잠재력은 충분함 [23]

- 지난해 약 150억 원의 매출을 기록했던 한라식품은 참치액 시장에서 약 40%의 비중 [24]
- 동원이 **15~20%의 시장 점유율**을 차지할 것이라 추정

제품 판매량은 다음과 같은 패턴일 거라 예상해 점유율 설정

900g < 500g 순 < 진 프리미엄 < 일반

동원참치액 월별 소비 패턴

동원참치액 순/진

1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
2.4	2.2	0.9	1.0	1.1	1.0	1.0	2.4	2.2	0.8	0.9	0.8

- 2024년 8-9월 (**추석 성수기**): 추석을 앞두고 선물세트 판매가 폭발적으로 증가하며 높은 판매량 예상
- 8월부터 사전예약 등으로 판매량이 급등
- 2024년 10월, 2025년 3월 (**명절 직후 비수기**): 명절 기간 동안 풀린 제품을 소진하는 시기로 예상 판매량 급락
- 2025년 1-2월 (**설 성수기**): 설을 대비한 선물세트 판매로 인해 다시 한 번 판매량이 폭발적으로 증가
- 그 외 기간: 명절을 제외한 기간에는 집밥 요리 빈도에 따라 비교적 안정적인 판매량을 유지

동원참치액 월별 소비 패턴

프리미엄 동원참치액 500g

1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
1.3	1.0	1.0	1.0	1.1	0.9	0.8	0.9	1.4	0.9	1.2	1.2

프리미엄 동원참치액 900g

1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
1.1	1.0	1.0	1.0	1.1	0.9	0.9	0.9	1.1	0.9	1.2	1.1

- 프리미엄은 선물 세트 비중이 적음
- 2024년 9월, 2025년 1월 (**명절 요리 성수기**): 명절 음식을 만들기 위해 구매하여 판매량이 증가
- 2024년 11월 (**김장/동절기 특수**): 김장 맛을 내거나, 겨울철 국물 요리를 위해 구매하는 수요가 증가
- 2024년 7-8월 (**여름 비수기**): 국물 요리를 잘 하지 않는 여름철에는 수요가 낮음
- 900g 제품: 특히 900g은 충성 고객과 고급 식당 위주일 것으로 예상, 12개월 간의 **판매량 변동이 가장 적고 안정적인 패턴으로 가중치 설정**

소화가 잘 되는 라떼 시장 규모



닐슨코리아의 2024년 조사에 따르면, 편의점에서 팔리는 액상커피는 연간 약 1조 500억 규모 [25]

- 편의점에서 구매할 수 있는 커피를 개당 약 2500원 책정하면, 매월 약 3000~3500만 개 정도 판매되는 것을 추측 가능
따라서 점유율이 낮은 커피라고 해도 매월 판매되는 개수는 상당히 클 것으로 예상
- 덴마크 소화가 잘 되는 우유로 만든 라떼는 출시 두 달 만에 누적 판매량 300만 개를 돌파 [26]

소화가 잘 되는 라떼 소비자 수

1600만명

타겟 고객 2133만명 중 75%

주 타겟 고객은 직장인 + 학생으로 설정

- 통계청에 따르면 대학생 수는 약 **235만명** [27]
- KOSIS 경제활동인구 총괄에 따르면 15~59세 취업자 수는 약 **2608만명** [28]

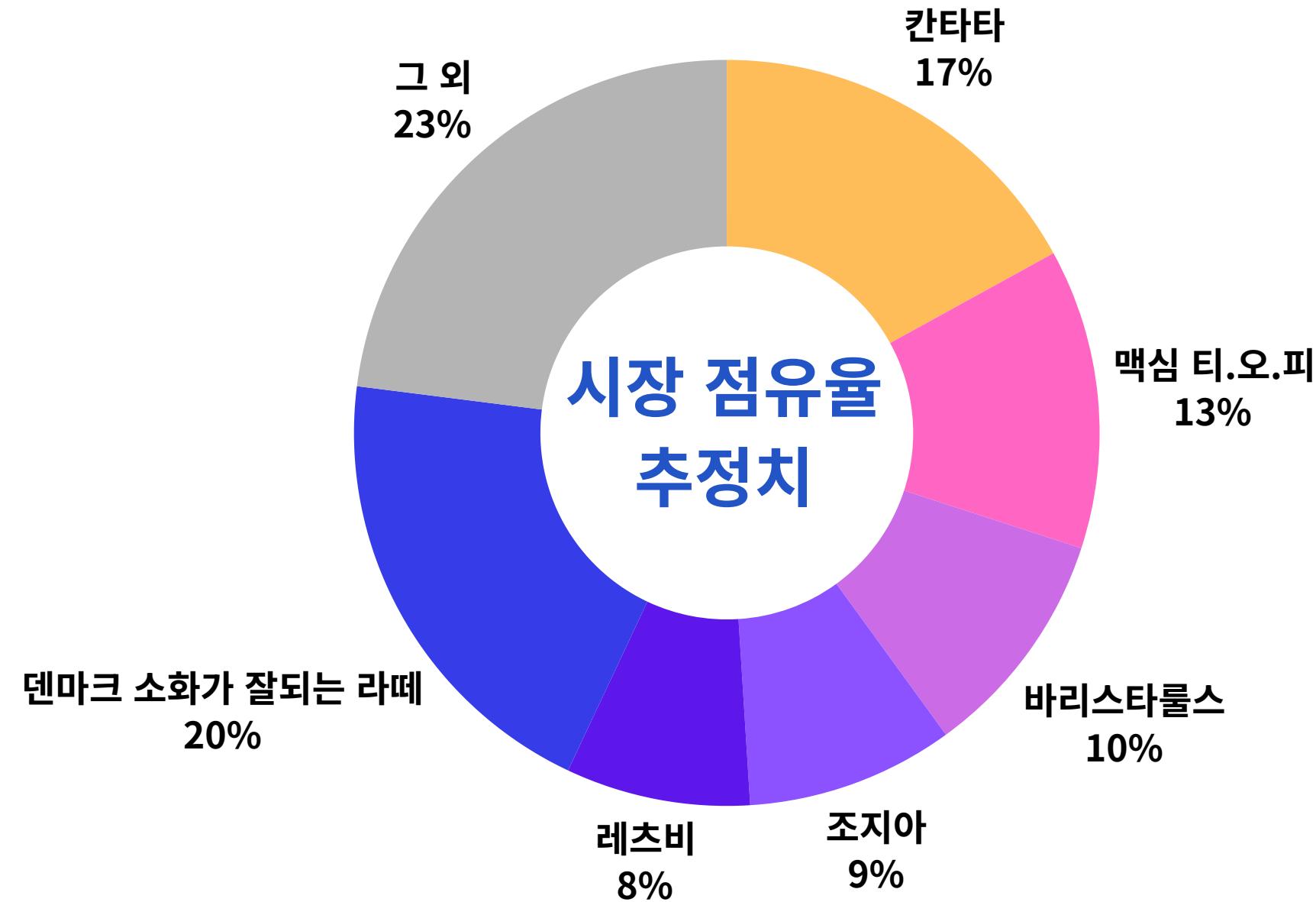
한국인의 연간 커피 소비량은 매우 높은데, 타겟 고객 중 마시는 비율은 전체의 75%로 산정

- $2843\text{만명} * 0.75 = \mathbf{2133\text{만명}}$

유당불내증은 전 세계 인구의 대략 75%를 차지,
동아시아인의 경우 더 높은 비율을 보임 [29][30]

- 아시아인은 75%보다 높은 비율이지만, 보수적으로 75% 정도로 산정
- $2133\text{만명} * 75\% = \mathbf{1600\text{만명}}$

소화가 잘 되는 라떼 소비자 수



락토프리 옵션이 실제 유효 수요 복원에 기여한다는 가정의 과학적 배경으로 사용

- 한국인은 유당불내증에 대해 대중적인 인지도가 있는 편이며, 이에 대한 논의도 활발함

일반적인 액상커피 시장 + 락토프리 시장이라는 특성을 결합한 시장의 점유율로 생각하는것이 효과적이라 판단

- 덴마크 우유 브랜드는 유당 zero 제품에 집중하고 있기에 **유당불내증 소비자들 중에 충성 고객이 상당히 존재할 것으로 추측**
- 카페라떼, 바닐라라떼를 합친 점유율을 약 20~25% 정도로 예상

소화가 잘 되는 라떼 월별 소비 패턴

덴마크 소화가 잘되는 우유로 만든 카페라떼/바닐라라떼

1월	2월	3월	4월	5월	6월	7월	8월	9월	10월	11월	12월
0.9	0.9	1.0	1.1	1.1	1.1	1.2	1.2	1.1	1.1	1.1	1.0

- 2024년 7-8월 (**여름**): 무더위로 인해 시원하고 달콤한 즉석 음료의 수요가 상승하여 비교적 높은 판매량
- 2025년 1월 ~ 2025년 2월 (**겨울** 비수기): 추운 날씨로 인해 차가운 컵커피 수요가 연중 최저점 [31]
- 2025년 4월, 6월 (시험 기간 특수): 대학가의 중간/기말고사 시즌에 학생들이 카페인과 당 충전을 위해 많이 구매하여 작은 고점을 형성할 가능성이 있음