

# 식음업장 메뉴 수요 예측 AI 오프라인 해커톤 솔루션 PPT

## LG Aimers



팀명: LGBestModel

최재우(최재우),  
오창석(오창석),  
MinsungBae(배민성),  
ICHINOSESHIKI(우호경),  
JooHyoung(채주형)

## Contents

식음업장 메뉴 수요 예측 AI 오프라인 해커톤

LG AIMERS

1. Executive Summary
2. 사용 데이터
3. 공통 코어 & Score- Aware 매핑
4. 공통 EDA 근거
5. Pipeline\_1 - 전처리 & 도메인
6. Pipeline\_1 - 시계열 FE & 학습
7. Tweedie p-CV & WindowCV
8. Pipeline\_2 - 전처리 & 도메인 보정
9. 추론 & 앙상블

# LG Aimers



# 1. Executive Summary

식음업장 메뉴 수요 예측 AI 오프라인 해커톤

LG AIMERS

최종 성과: Public [0.56223], Private [0.55326]

## 핵심 전략 4가지

1. Score-Aware 설계
2. ROOM 0-day 보정,
3. p\*앵커 듀얼 앙상블
4. 메디안 결합

한 줄: 산식에 직접 매핑된 전처리 + 서로 다른 도메인 튜닝  
→ 강력한 성능

## 2. 사용 데이터

식음업장 메뉴 수요 예측 AI 오프라인 해커톤

LG AIMERS

train.csv			price.csv			room_type.csv			TRAIN_group.csv			TRAIN_room.csv			TRAIN_hwadam.csv		
Views			Grid view			Hide fields			Filter			Group			Sort		
<input type="checkbox"/>	영업일자	영업장명_메뉴명	매출수량														
1	2023-01-01	느티나무 셀프BBQ 1인	0														
2	2023-01-02	느티나무 셀프BBQ															
3	2023-01-03	느티나무 셀프BBQ															
4	2023-01-04	느티나무 셀프BBQ															
5	2023-01-05	느티나무 셀프BBQ															
6	2023-01-06	느티나무 셀프BBQ															
7	2023-01-07	느티나무 셀프BBQ															
8	2023-01-08	느티나무 셀프BBQ															
9	2023-01-09	느티나무 셀프BBQ															
10	2023-01-10	느티나무 셀프BBQ															

train.csv			price.csv			room_type.csv			TRAIN_group.csv			TRAIN_room.csv			TRAIN_hwadam.csv			TRAIN_ski.csv			TRAIN_weather		
Views			Grid view			Hide fields			Filter			Group			Sort								
<input type="checkbox"/>	일시	평균기온(°C)	최고기온(°C)																				
1	2023-01-01	-1.7																					
2	2023-01-02	-4.1																					
3	2023-01-03	-6.2																					
4	2023-01-04	-3.8																					
5	2023-01-05	-5.3																					
6	2023-01-06	-2.7																					
7	2023-01-07	0.9																					
8	2023-01-08	-1.9																					
9	2023-01-09	2.7																					
10	2023-01-10	-0.2																					

train.csv			price.csv			room_type.csv			TRAIN_group.csv			TRAIN_room.csv			TRAIN_hwadam.csv			TRAIN_ski.csv			TRAIN_weather		
Views			Grid view			Hide fields			Filter			Group			Sort								
<input type="checkbox"/>	영업일자	느티나무 셀프BBQ	담하			라그로타			미라시아														
1	2023-01-01	0	0			0																	
2	2023-01-02	0	6			0																	
3	2023-01-03	4	4			0																	
4	2023-01-04	0	2			0																	
5	2023-01-05	4	2			0																	
6	2023-01-06	0	5			1																	
7	2023-01-07	0	4			0																	
8	2023-01-08	4	0			0																	
9	2023-01-09	0	4			2																	
10	2023-01-10	0	1			0																	

사용 데이터: train.csv, TRAIN\_weather, TRAIN\_group  
(+ P1: TRAIN\_room, room\_type; P2: TEST\_room)

### 3. 공통 코어 & Score - Aware 매핑

Pipeline\_1 & Pipeline\_2

LG AIMERS



**공통 파이프라인: 안전 전처리(평일↔공휴일 보간, 급등 클램프)**

- 시계열 FE
- LGBM Tweedie p-앙상블
- 앵커(K=7)
- Metric-aware 후처리

**산식 매핑:**

A(SMAPE) ↔ 급등 클램프·라운딩

B/C(MAE·RMSE/ $\bar{y}$ ) ↔ 규모 신호·min-1·ROOM 보정

D( $r^2$ ) ↔ 앵커·계절/이벤트 신호

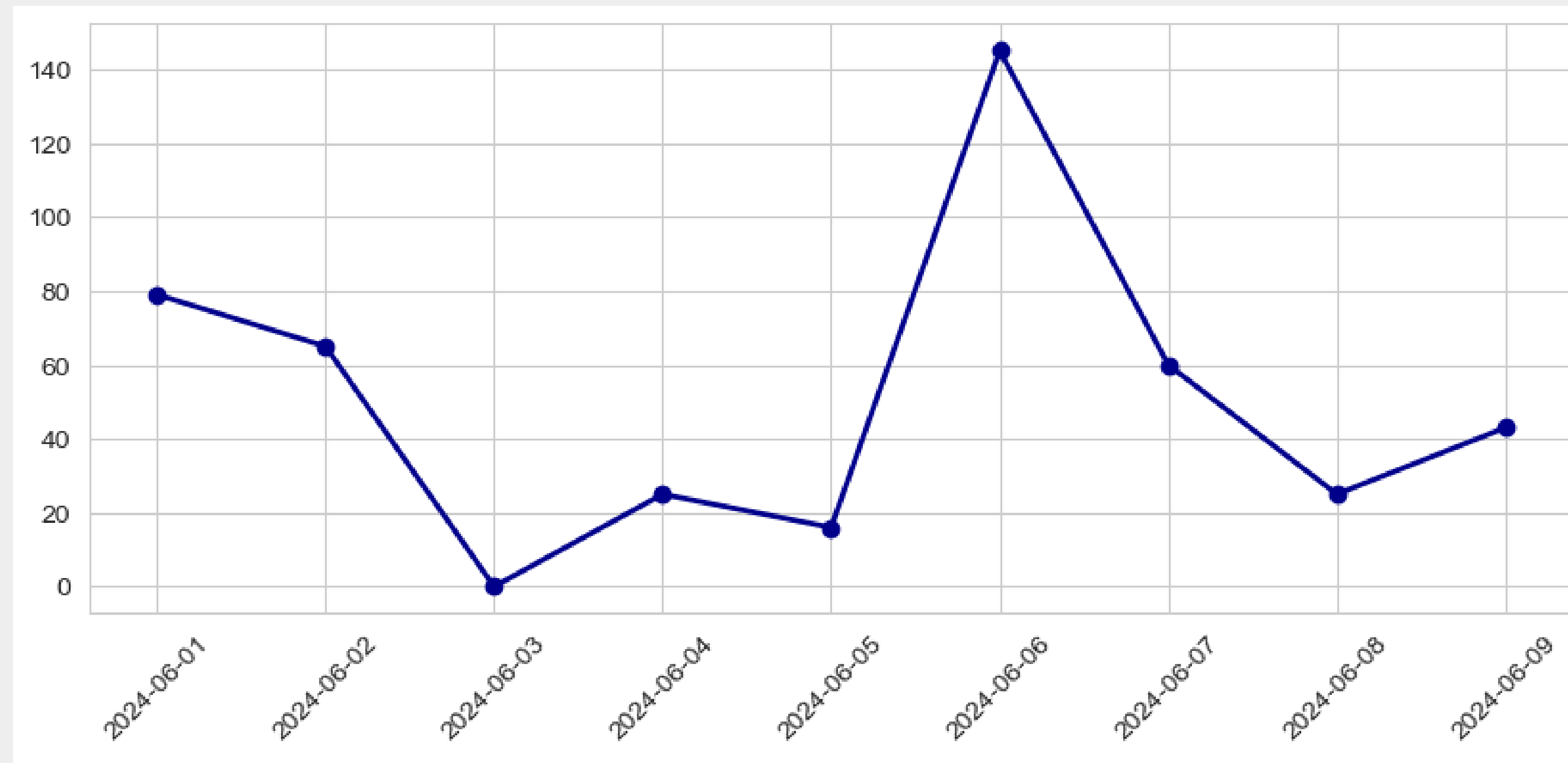
**같은 엔진, 다른 튜닝으로 다양성 확보**

## 4. 공통 EDA 근거

EDA 기반 전략 및 근거

LG AIMERS

위에 화담숲카페\_현미뱅크크림, 왼쪽에 매출수량



**패턴: 시즌/요일/공휴일 차이 뚜렷(캘린더 히트맵, 요일 박스)**

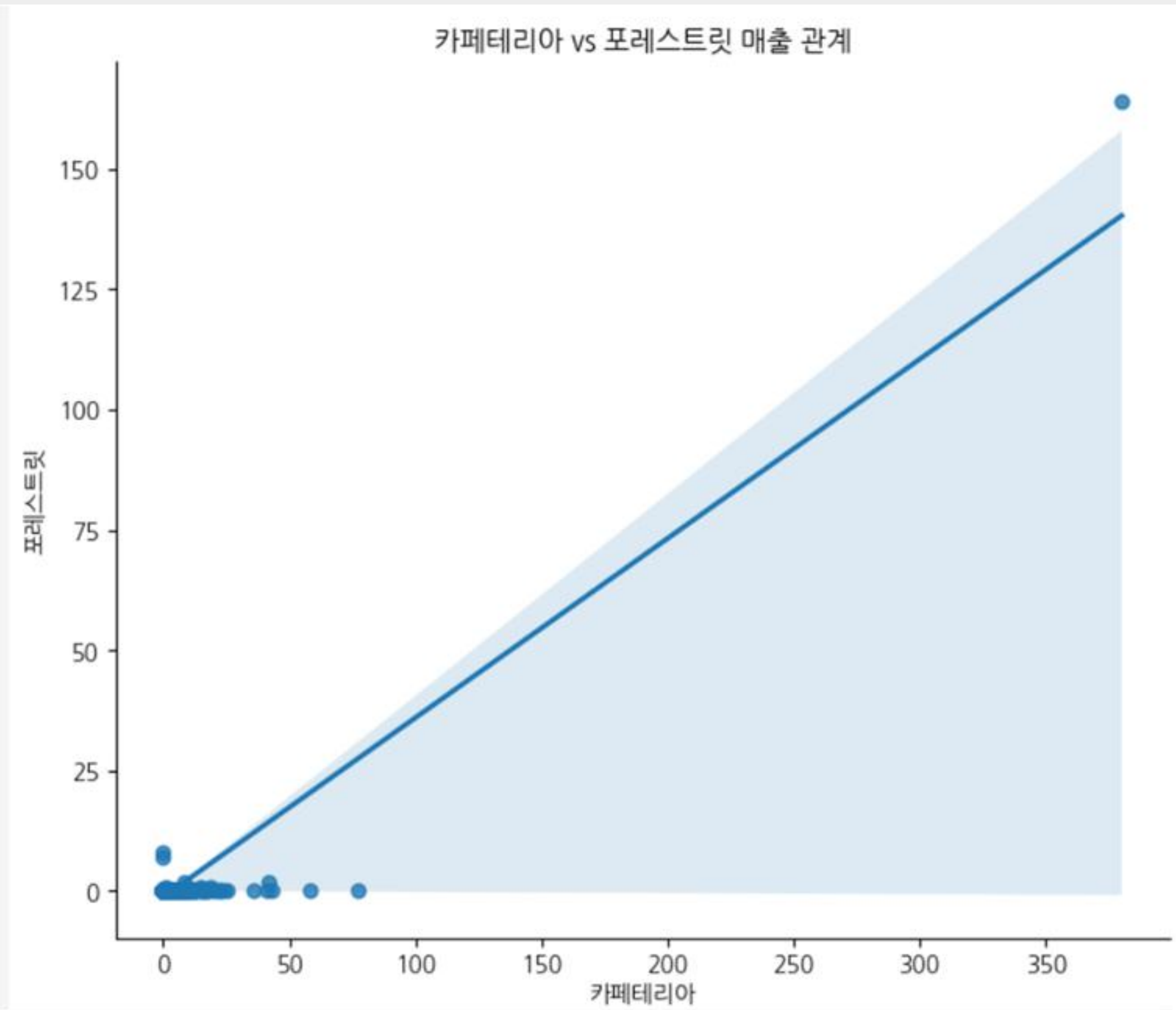
**보간 효과: 평일 공휴일의 과도한 급등치 완화 → sMAPE 꼬리 감소**

**변동 관리: 클램프·미래 캘린더(H1~H7)로 분산 안정화**

## 5. Pipeline\_1 - 전처리 & 도메인

EDA 기반 전략 및 근거

LG AIMERS



그룹 이벤트: 카페테리아+포레스트릿 합 → **is\_forest\_event** 플래그

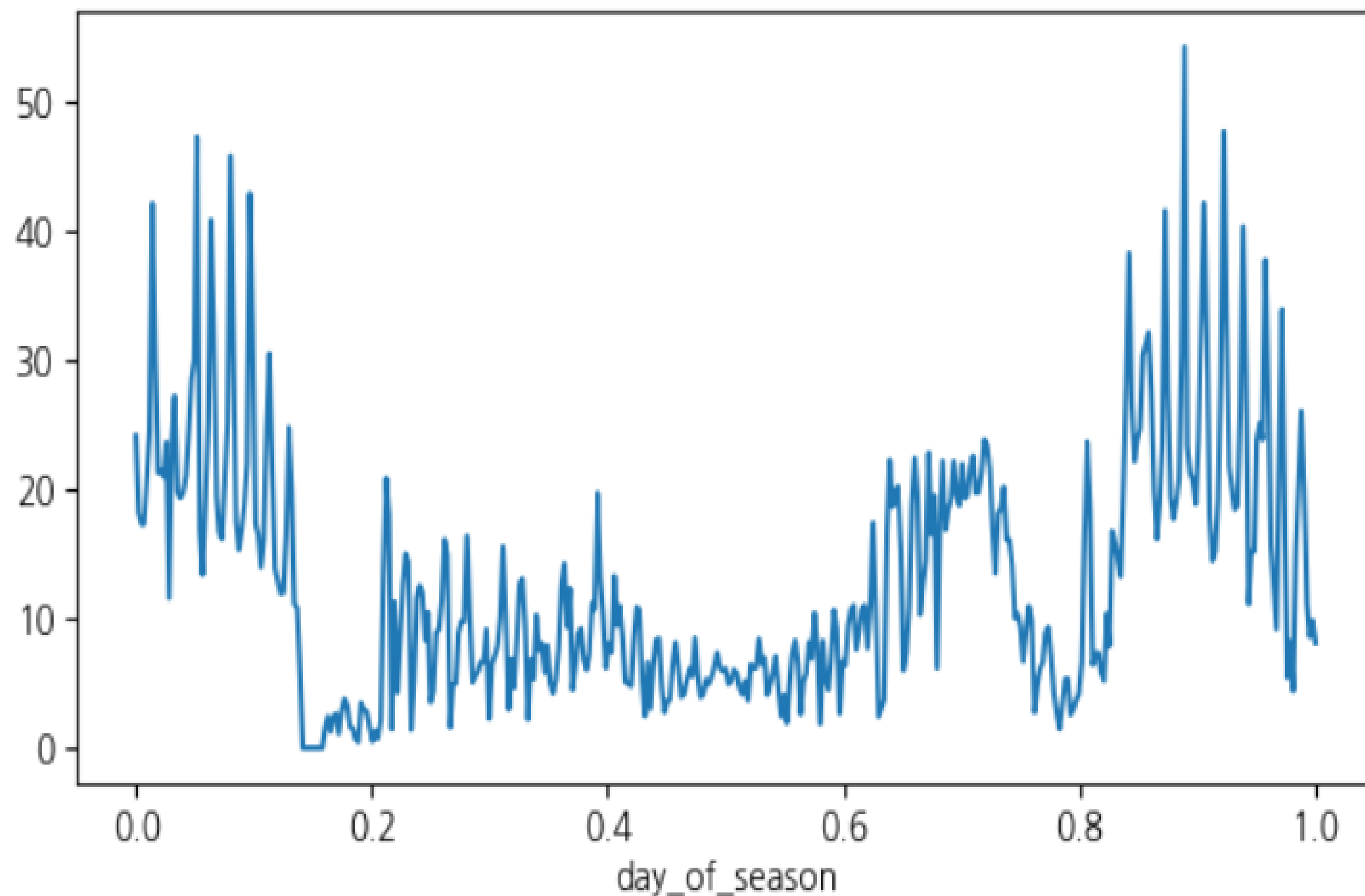
날씨: 일별 최고/평균기온 집계

## 5. Pipeline\_1 - 전처리 & 도메인

EDA 기반 전략 및 근거

LG AIMERS

시즌 상대 위치별 매출



겨울 시즌: 플래그 + **\*\*시즌 상대위치(0~1)\*\***로 초/중/후 효과 반영

HOL\_WEEKDAY\_SET: 평일 공휴일 분리

급등 완화: 단체/특정키(EXTRA\_GROUP\_KEYS) **타킷 안정화**

규모 신호: room\_type\*기준인원 → 총방문객수

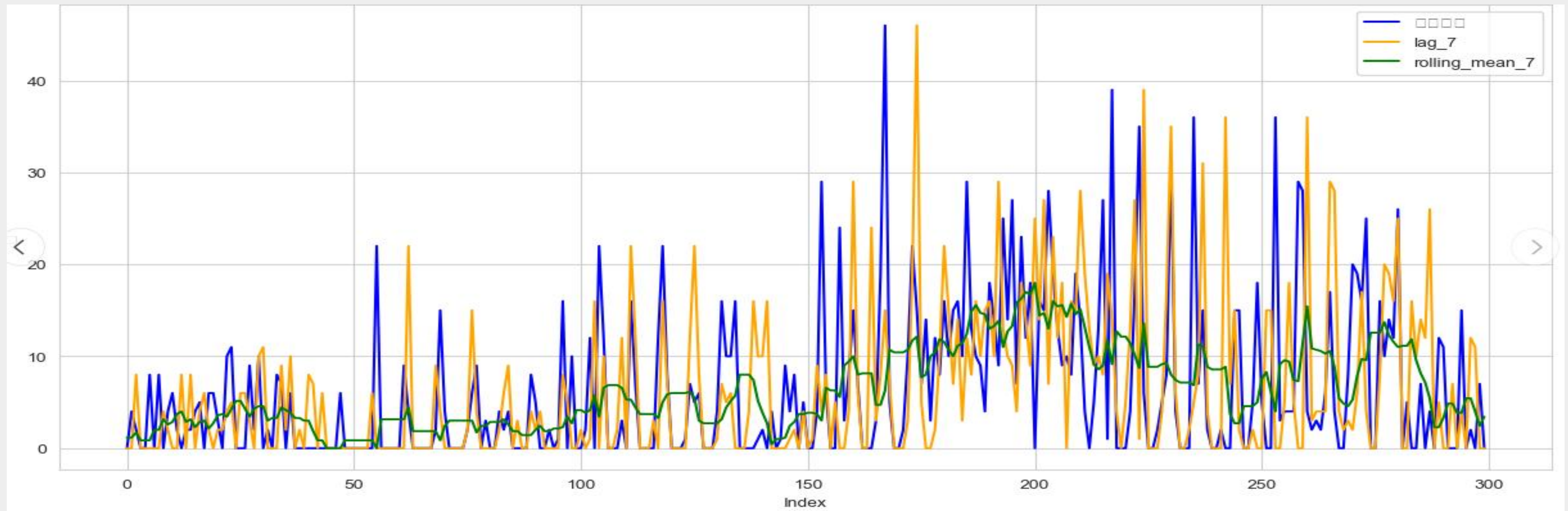
장기 미판매: ≥14일 0-블록 플래그



## 6. Pipeline\_2 - 시계열 FE & 학습

추세 및 변동 탐지

LG AIMERS



**FE: Lag(1~21), Rolling, Slope(7/14), EWM(7), Fourier(dow\_sin/cos), 달력/미래 DOW**

**KEY 인코딩: 영업장명\_메뉴명 카테고리**

**사용 피처: 총 58개(현재 세팅)**

**메시지: 패턴·추세·변동을 균형 있게 포착**

## 7. Tweedie p-CV & WindowCV

검증 성능 평가 및 방향성 확립

LG AIMERS

**데이터 특성: 0 다수 + 양수 연속 Tweedie**

**5-Fold 결과:**

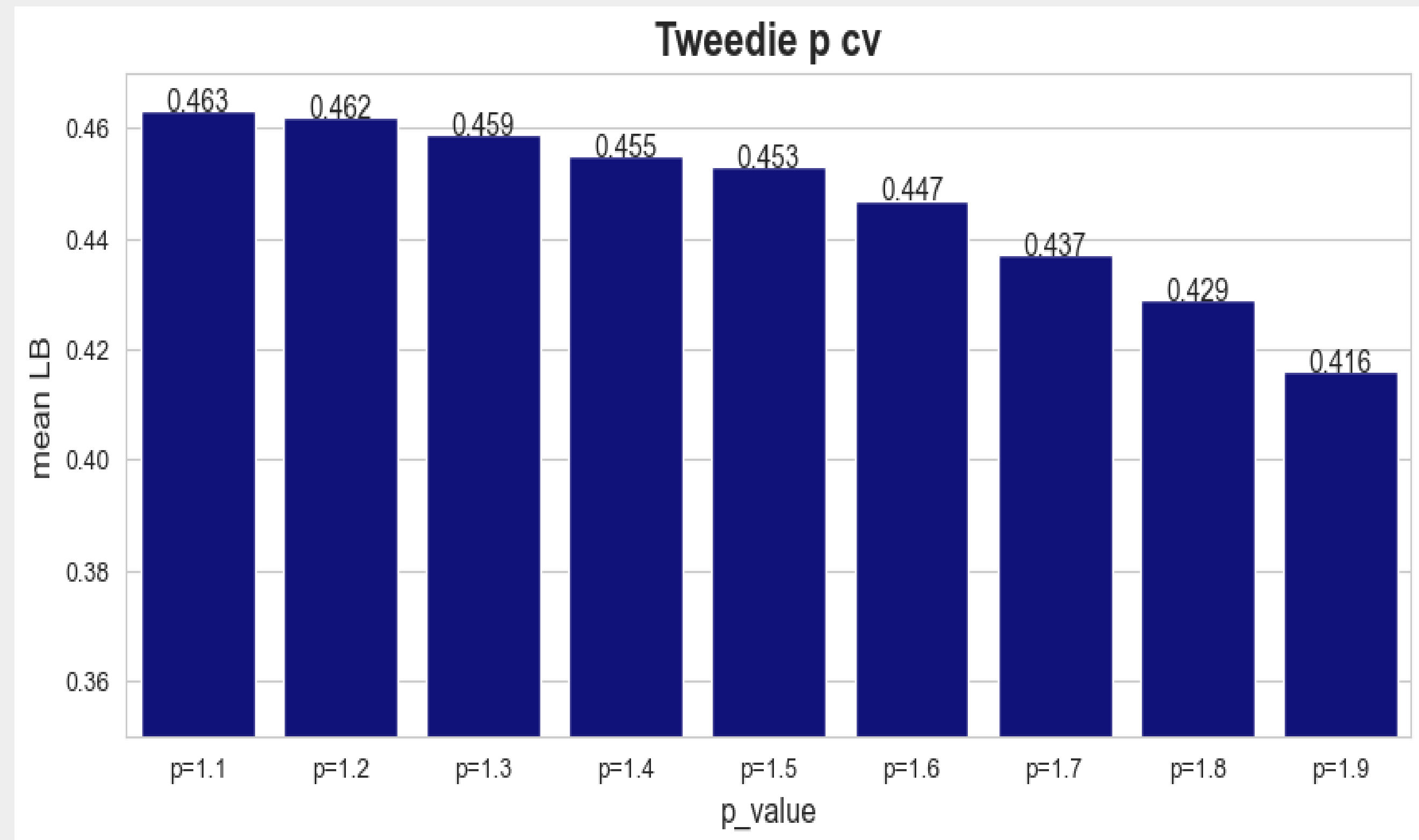
**p=1.1: mean LB 0.46346**

**p=1.2: 0.46265**

**p=1.3: 0.45991 (이후 하락 추세)**

**검증 설계: WindowCV(입력 35일 → 예측 7일, stride=7, folds=5)**

**결론:  $p \in [1.1, 1.3]$  우세 → p-**양상블로 분포 불확실성 완화****



## 8. Pipeline\_2 - 전처리 & 도메인보정

EDA 기반 전처리

LG AIMERS

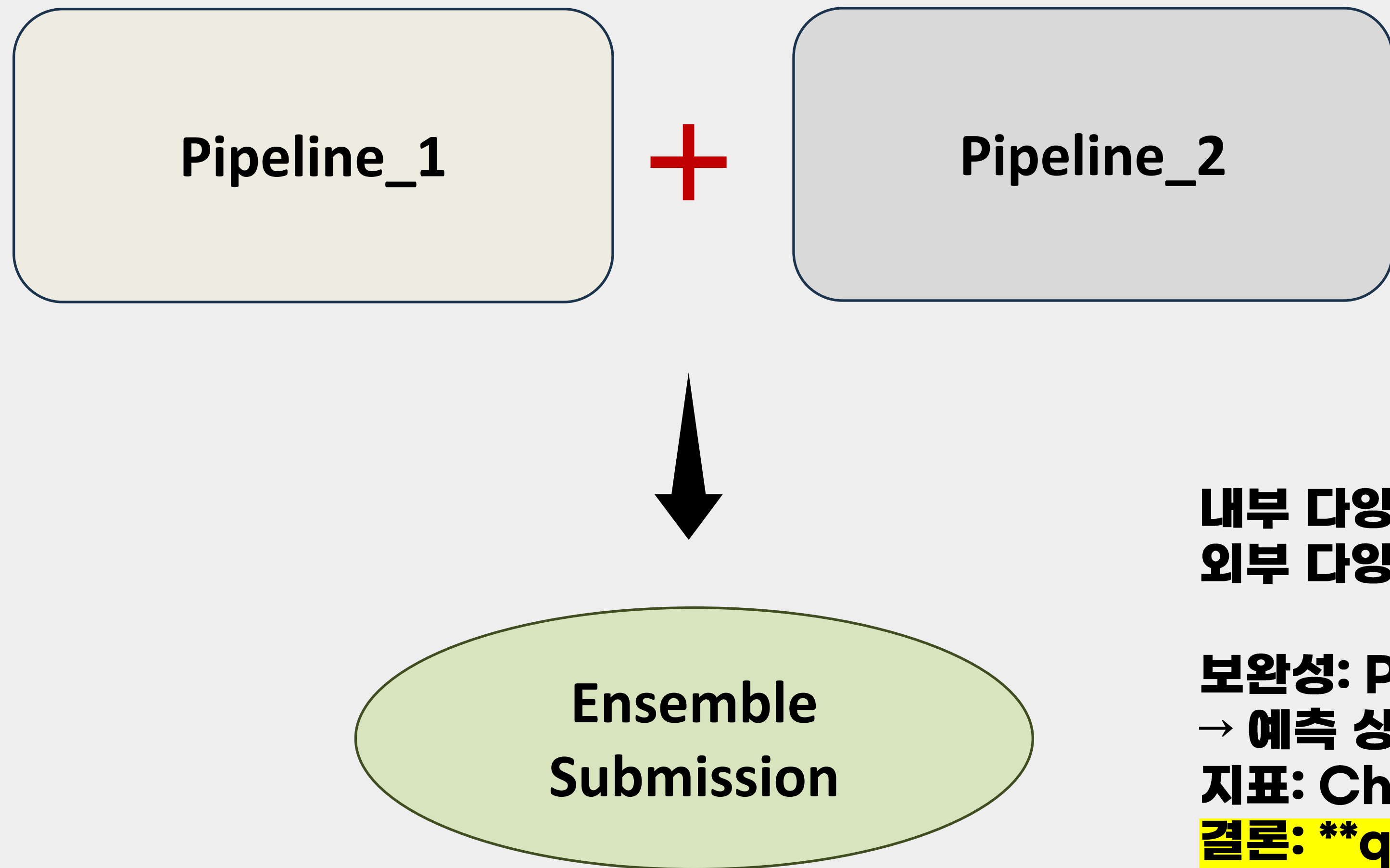
1. ROOM 0-day 보정: 객실총합=0일  
→ **같은 주중/주말 과거 정상일 중앙값** 대체(Train/Test)
2. 마지막 클램프 강화
3. Hwadam 피쳐: 화담숲주막/카페에만 게이팅 적용

제로/스파이크 억제로 A/B/C/D 전반 안정화

## 9. 추론 및 Ensemble

최종 예측 및 앙상블 전략

LG AIMERS



**내부 다양성: p-앙상블 \* 앵커(K=7)**

**외부 다양성: P1 \* P2 메디안 앙상블(+ min-1)**

**보완성: P1(상향·패턴 민감) vs P2(중앙경향·보정 강화)**

→ **예측 상관 ↓**

**지표: Changed-cell ratio [x.xx%] (자리표시자)**

**결론: \*\*q50(메디안)\*\*으로 꼬리/이상치 방어**