

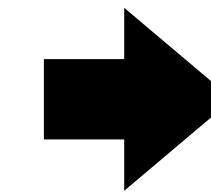
# C2C 중고거래 가격 추천 시스템

김민지, 박지원, 배성환, 이하준



# 개발 배경

물가 상승에 따른 소비자들의 지출 감소  
효율적인 구매를 원하는 소비자들의 중고 거래 사용량 증가



중고 거래 시장 급성장

## 고물가에 중고거래 시장 급성장

기사입력 2024-02-04 14:03:52



서울 현대백화점 신촌점 유플렉스 내 중고품 전문관 2022년 개장 당시 모습./사진=대한경제 DB

[대한경제=오진주 기자] 고물가에 리커머스(Re-commerce) 시장이 활발해지고 있다.

4일 한국인터넷진흥원에 따르면 오는 2025년 국내 중고 거래 시장은 43조원 규모로 커질 것으로 예상된다. 중고 거래 시장 규모는 지난 2008년 4조원에서 2021년 24조원으로 급성장했다.

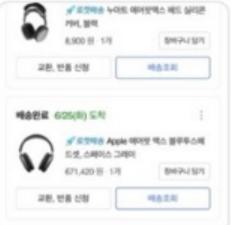
출처 : 대한경제



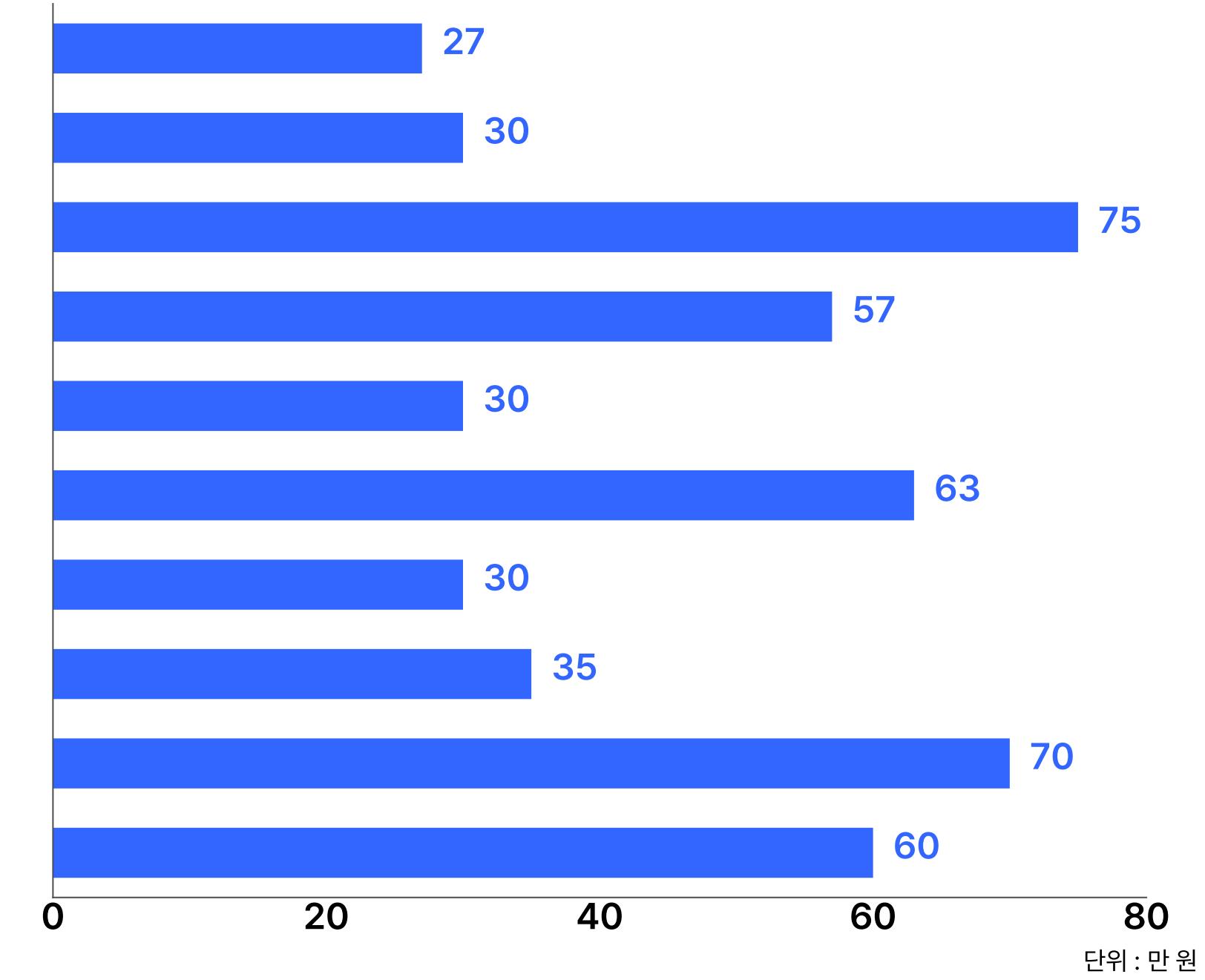
출처 : 연합뉴스

# 개발 배경

같은 상품임에도 가격대가 제각각이라, 사용자들은 합리적인 가격 파악이 어려움

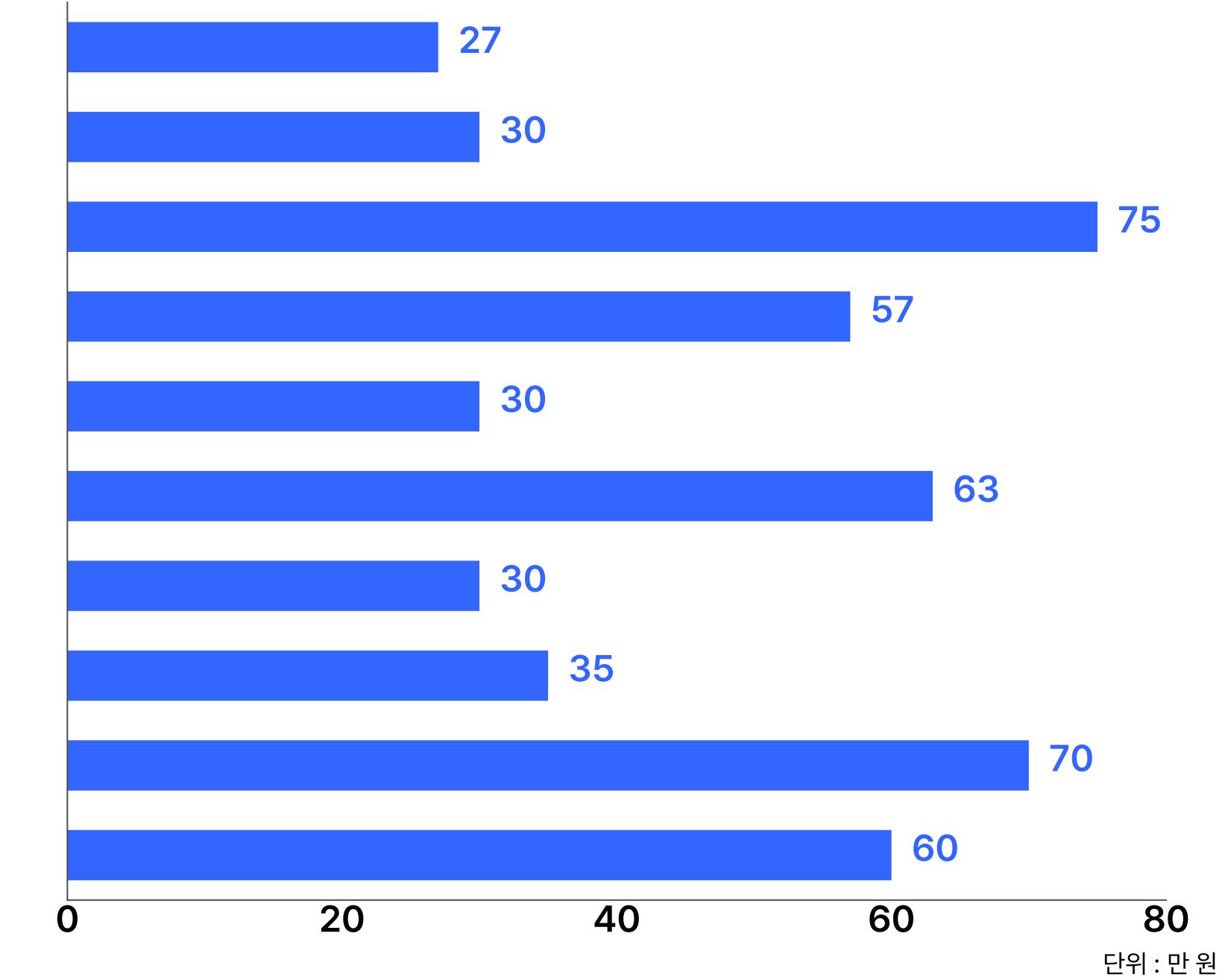
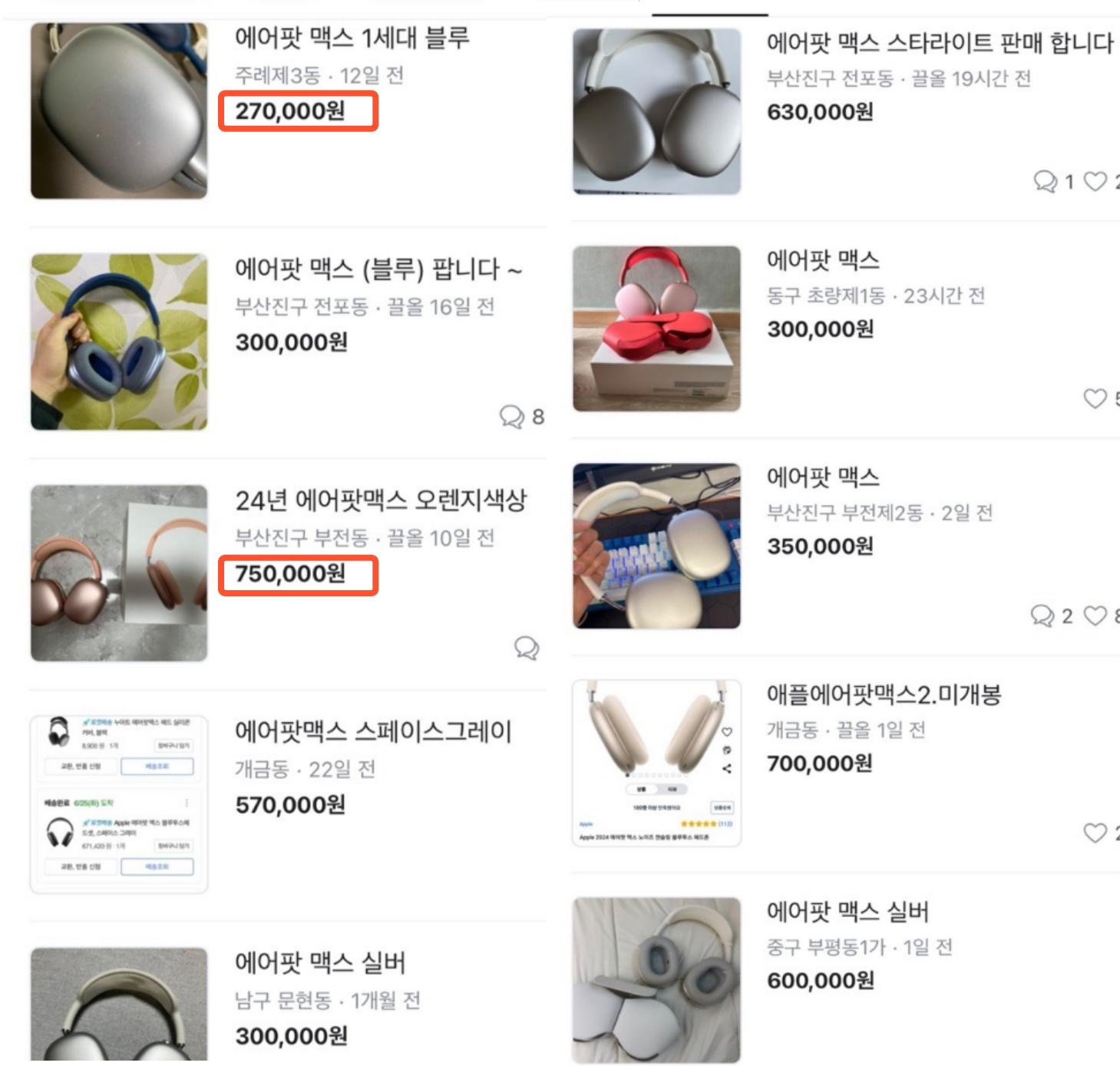
	에어팟 맥스 1세대 블루 주례제3동 · 12일 전 <b>270,000원</b>		에어팟 맥스 스타라이트 판매 합니다 부산진구 전포동 · 끌을 19시간 전 <b>630,000원</b>
	에어팟 맥스 (블루) 팝니다 ~ 부산진구 전포동 · 끌을 16일 전 <b>300,000원</b>		에어팟 맥스 동구 초량제1동 · 끌을 23시간 전 <b>300,000원</b>
	24년 에어팟맥스 오렌지색상 부산진구 부전동 · 끌을 10일 전 <b>750,000원</b>		에어팟 맥스 부산진구 부전제2동 · 2일 전 <b>350,000원</b>
	에어팟맥스 스페이스그레이 개금동 · 22일 전 <b>570,000원</b>		애플에어팟맥스2.미개봉 개금동 · 끌을 1일 전 <b>700,000원</b>
	에어팟 맥스 실버 남구 문현동 · 1개월 전 <b>300,000원</b>		에어팟 맥스 실버 중구 부평동1가 · 1일 전 <b>600,000원</b>

출처 : 당근마켓



# 개발 배경

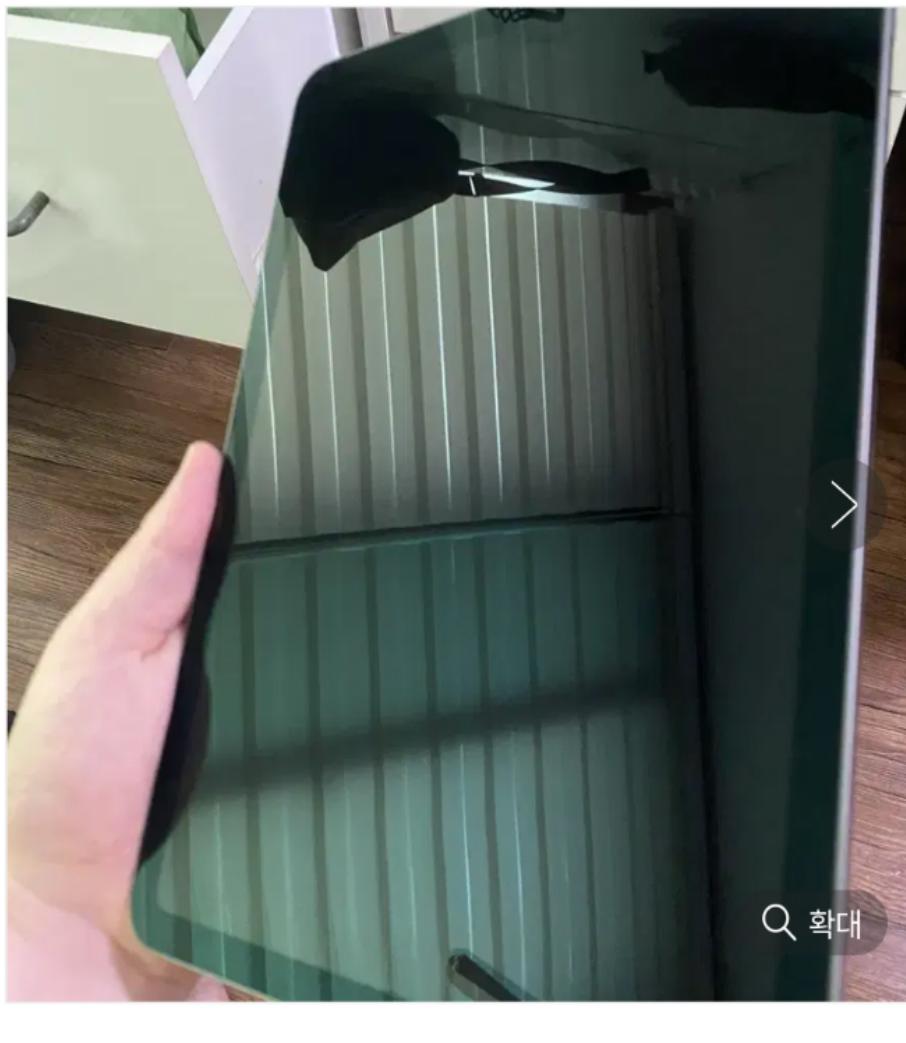
같은 상품임에도 가격대가 제각각이라, 사용자들은 합리적인 가격 파악이 어려움



# 사용자에게 합리적인 가격 범위 추천

# 데이터셋 주요 속성

## 중고 물품 데이터 수집



출처 : 번개장터

아이패드 에어 4 64 gb

545,000 원

찜 10 | 뷰 248 | 10시간 전

- 상품상태 사용감 없음
- 배송비 일반 4,000원
- 직거래지역 서울특별시 강남구 논현1동

### 데이터 특성

상품 상태, 좋아요 수, 조회 수, 등록 기간

가격 예측에 필요한 데이터

상태 등급	상태 설명
1	고장/파손
2	사용감 많음
3	사용감 적음
4	사용감 없음
5	새 상품

13개의 카테고리에서 데이터를 수집 » 각 카테고리당 500개~5000개의 데이터 확보  
약 4만 개 이상의 데이터를 수집

# 데이터 전처리 결과

## 데이터 정제 기준

- ① 정가를 초과하는 가격 제거: 중고 거래에서 정가보다 높은 가격은 거래 가능성이 낮아 제거
- ② 지나치게 낮은 가격 제거: 현실적이지 않은 매우 낮은 가격 제거

항목	전처리 전 데이터 개수	전처리 후 데이터 개수	데이터 로스율
아이폰15 프로	1368	1173	14.25%
아이폰15	586	496	15.36%
에어팟 맥스	2189	1775	18.91%
에어팟 프로	4722	3909	17.22%

# 가격 예측을 위한 특성 중요도 평가

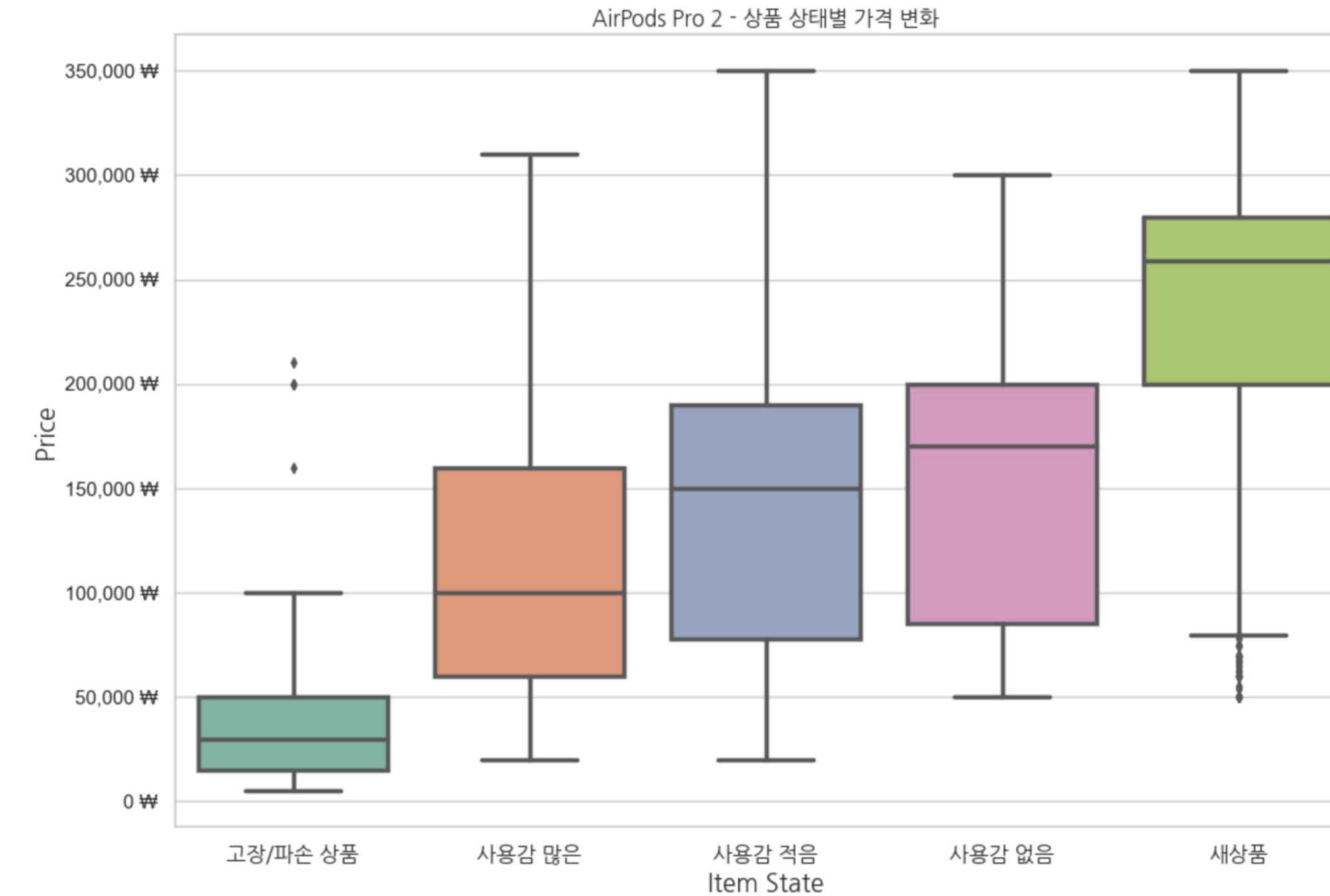
가격 예측에 있어 어떤 특성이 가장 큰 영향을 미칠까?

여러 가지 특성 중 상품 상태가 가격에 가장 큰 영향을 미칠 것이라고 가설 세움

# 가격 예측을 위한 특성 중요도 평가

가격 예측에 있어 어떤 특성이 가장 큰 영향을 미칠까?

여러 가지 특성 중 상품 상태가 가격에 가장 큰 영향을 미칠 것이라고 가설 세움



상품 상태별 가격 변화를 분석한 결과, 상태에 따라 뚜렷한 가격 차이가 나타남

# 가격 예측을 위한 특성 중요도 평가

## ANOVA(Analysis of Variance) 분석

여러 그룹 간의 평균 차이를 검정하여, 각 특성이 가격에 미치는 영향이 통계적으로 유의미한지를 확인하는 데 효과적

“ 상품 상태가 다른 특성에 비해 가격에 얼마나 큰 영향을 미치는지 검증 ”

- F-통계량

F 값이 클수록 해당 특성이 가격에 미치는 영향이 크다는 의미

- P-값

P-값이 0.05보다 작으면, 해당 특성이 가격과 통계적으로 유의미한 연관성을 가진다는 의미

# 가격 예측을 위한 특성 중요도 평가

- 상품 상태

F-Statistic (상품상태) : 320.21600546340164

p-value (상품상태) : 0.00798109408109112

$\rightarrow$  p-value < 0.05이므로 유의미함

- 조회수

F-Statistic (조회수) : 4.5205588987112835

p-value (조회수) : 0.01093407934987577

- 좋아요 수

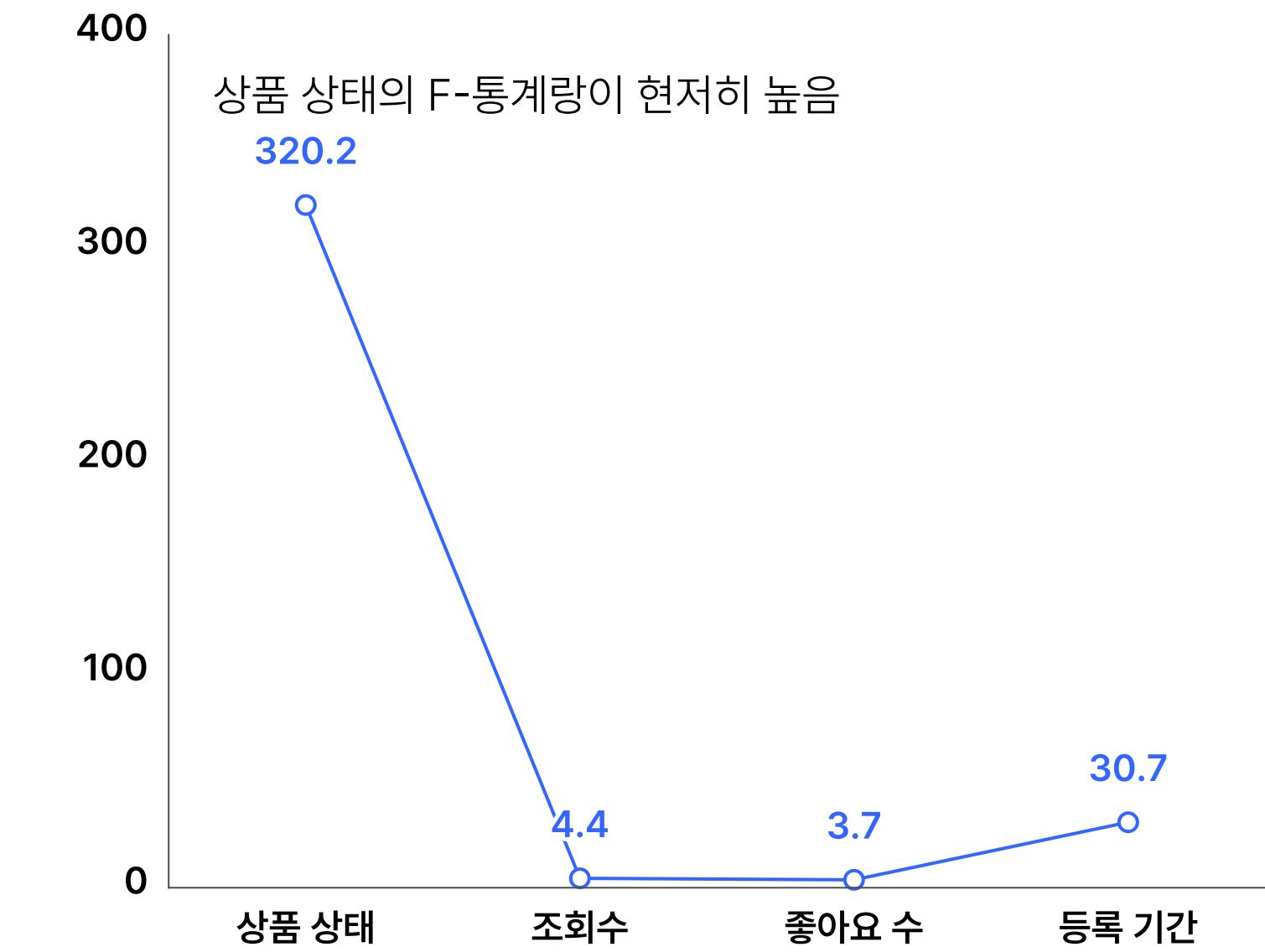
F-Statistic (찜 수) : 3.6565371230269283

p-value (찜 수) : 0.025901120381695796

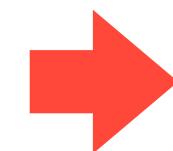
- 등록기간

F-Statistic (등록 기간) : 30.682241857503385

p-value (등록 기간) : 0.00586031213612045



상품 상태가 가격에 큰 영향을 미친다는 점을 통계적으로 입증



상품 상태를 기준으로 가격을 예측하여 사용자에게 추천

# 가격 예측을 위한 모델 선정

## Random Forest

다수의 결정 트리를 결합하여 최종 예측을 만드는 앙상블 모델

## XGBoost

이전 모델의 예측 오차를 보완하며 성능을 개선하는 부스팅 알고리즘

## Gradient Boosting

점진적으로 성능을 개선하기 위해 순차적으로 학습하는 부스팅 기법

# 가격 예측을 위한 모델 선정

## Random Forest

- **장점** : 비선형 데이터에 적합하며 과적합 방지 효과가 큼
- **단점** : 복잡한 데이터에서 처리 속도가 느려질 수 있음

## XGBoost

- **장점** : 복잡한 데이터에서 뛰어난 성능을 보이며, 학습 속도가 빠름
- **단점** : 많은 계산 자원을 필요로 하며, 과적합 가능성 있음

## Gradient Boosting

- **장점** : 작은 데이터셋과 특정 패턴 학습에 유리함
- **단점** : 학습 속도가 느리고, 파라미터 조정이 까다로울 수 있음

# 가격 예측을 위한 모델 선정 이유

- ① Random Forest: 상품 상태와 같이 중요한 변수가 있을 때 빠르고 안정적인 성능을 기대할 수 있음
- ② XGBoost: 고가 전자제품 등 복잡한 데이터셋에서 정확한 예측을 기대할 수 있음
- ③ Gradient Boosting: 상대적으로 적은 데이터에서 빠른 트렌드 파악에 유리

# 가격 예측을 위한 모델 선정 이유

- ① Random Forest: 상품 상태와 같이 중요한 변수가 있을 때 빠르고 안정적인 성능을 기대할 수 있음
- ② XGBoost: 고가 전자제품 등 복잡한 데이터셋에서 정확한 예측을 기대할 수 있음
- ③ Gradient Boosting: 상대적으로 적은 데이터에서 빠른 트렌드 파악에 유리

## 모델 성능 평가

### R<sup>2</sup> (결정계수)

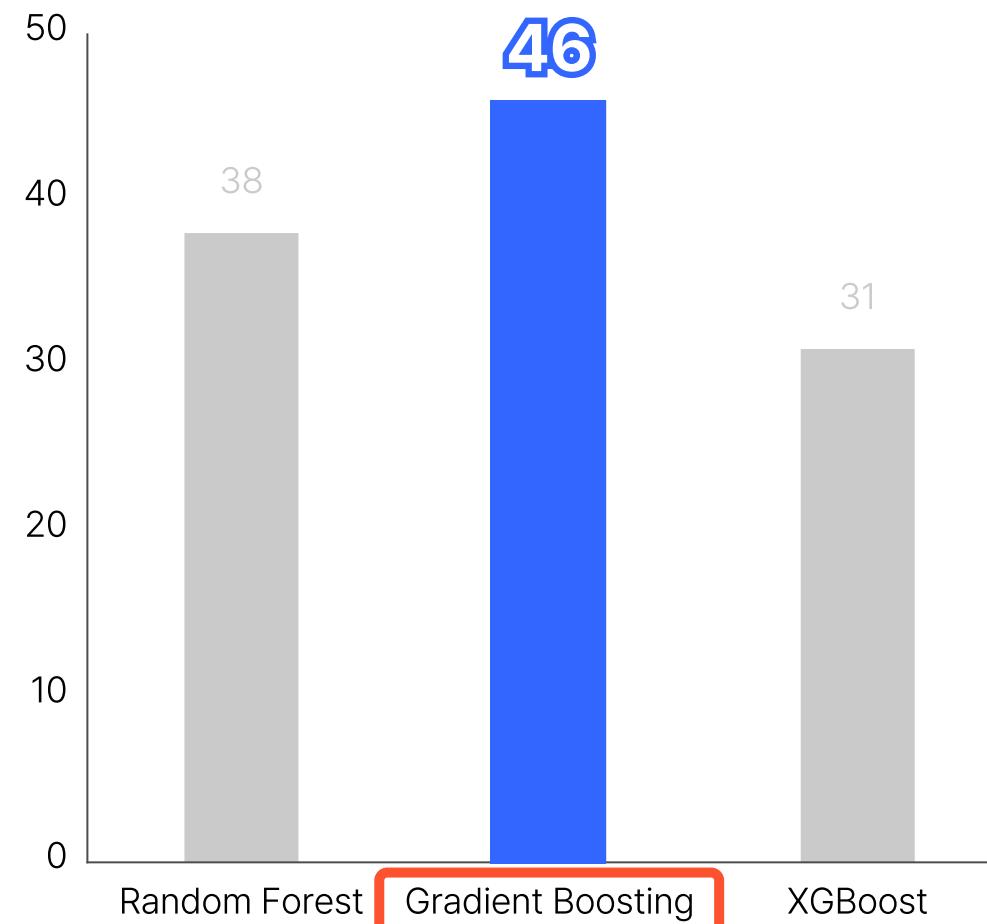
- 모델이 가격을 얼마나 잘 설명하는지를 나타내는 지표
- 값이 1에 가까울수록 예측력이 높음

# 모델 성능 평가

데이터셋 특성에 따라 최적 성능의 모델이 달라짐을 확인

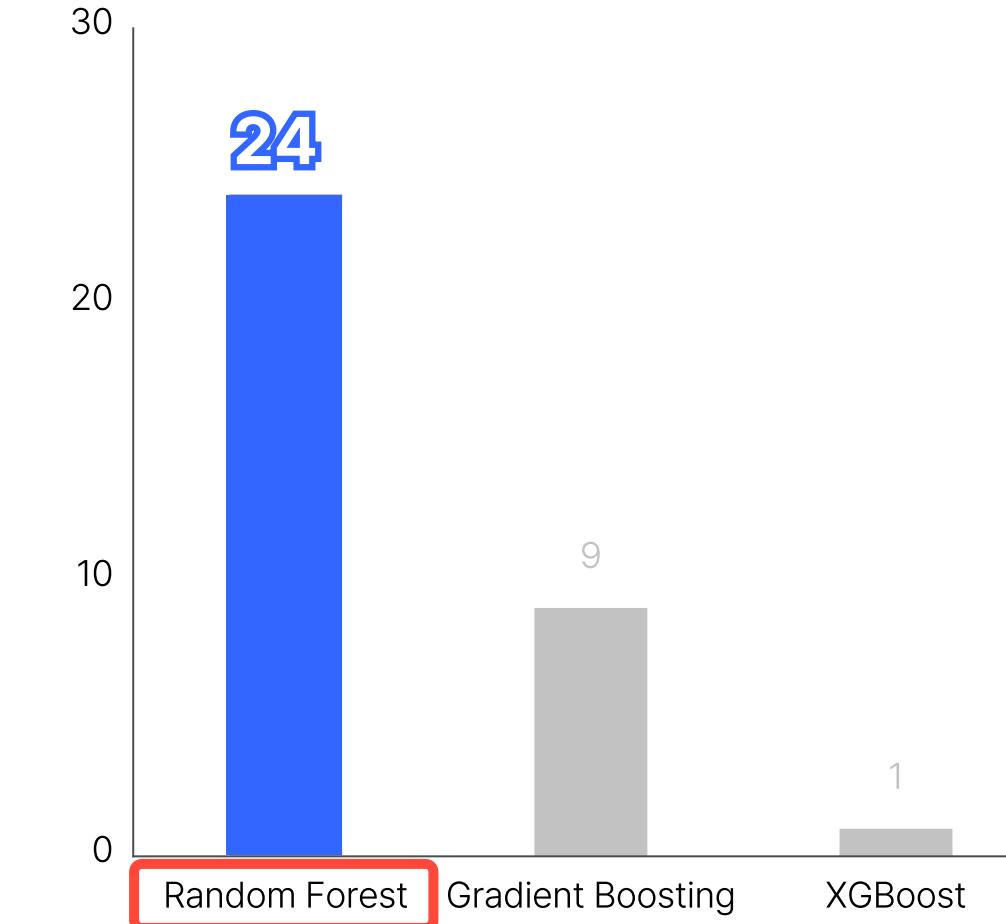
AirPods Max

R <sup>2</sup> (모델 성능 평가 지표)	
Random Forest	0.378069
Gradient Boosting	0.458389
XGBoost	0.305196



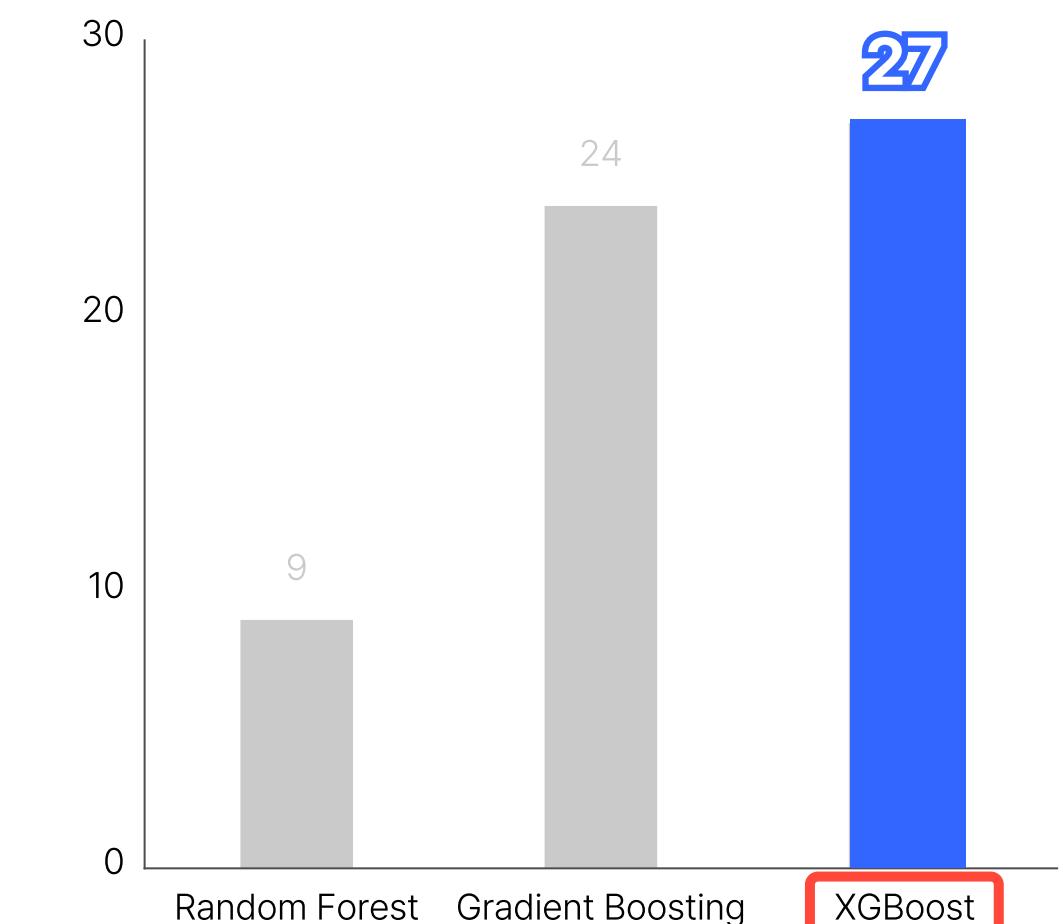
AirPods Pro2

R <sup>2</sup>	
Random Forest	0.243323
Gradient Boosting	0.086759
XGBoost	0.010367



Iphone 15

R <sup>2</sup>	
Random Forest	0.086421
Gradient Boosting	0.235246
XGBoost	0.272450



# 모델 성능 평가

AirPods Max

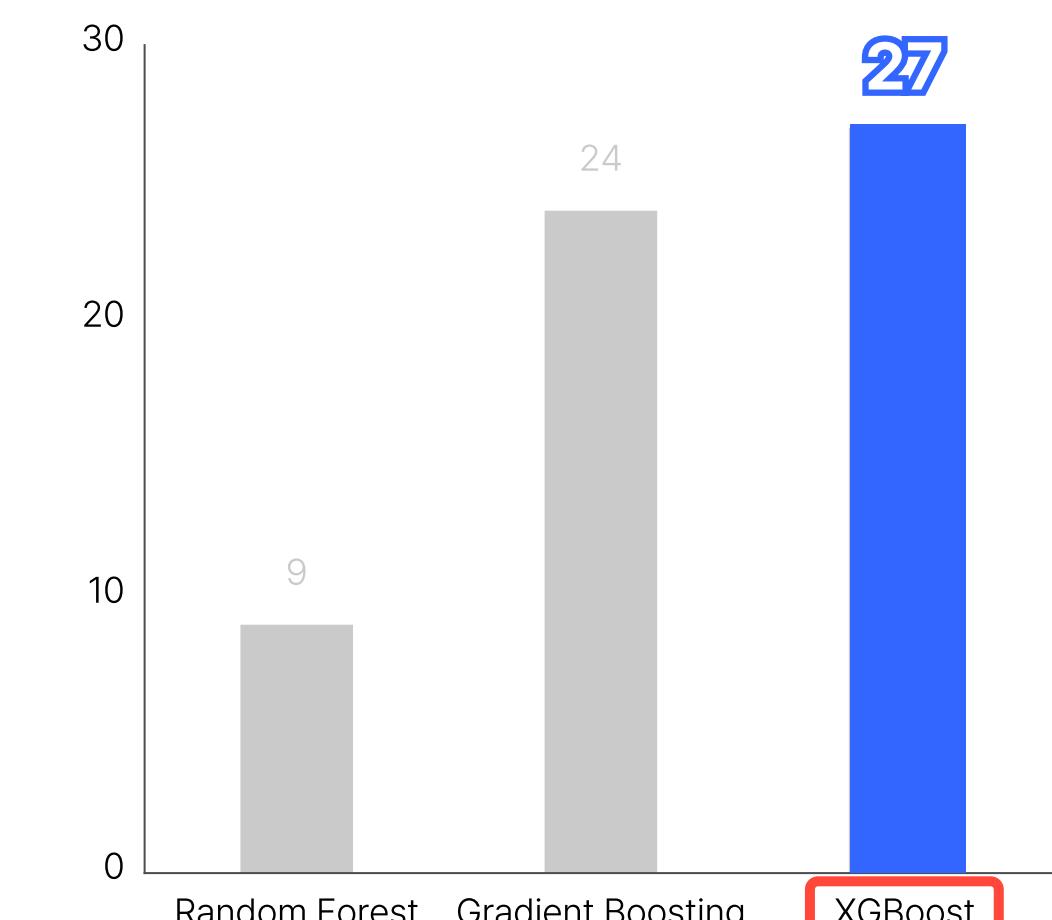
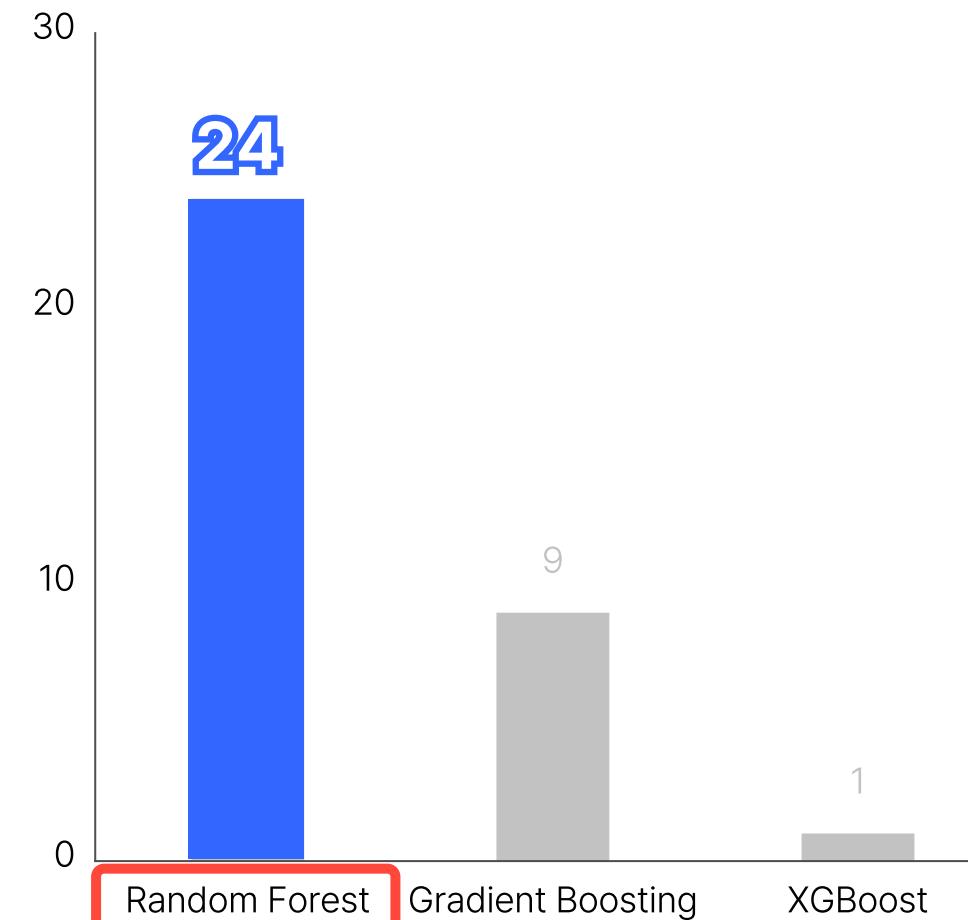
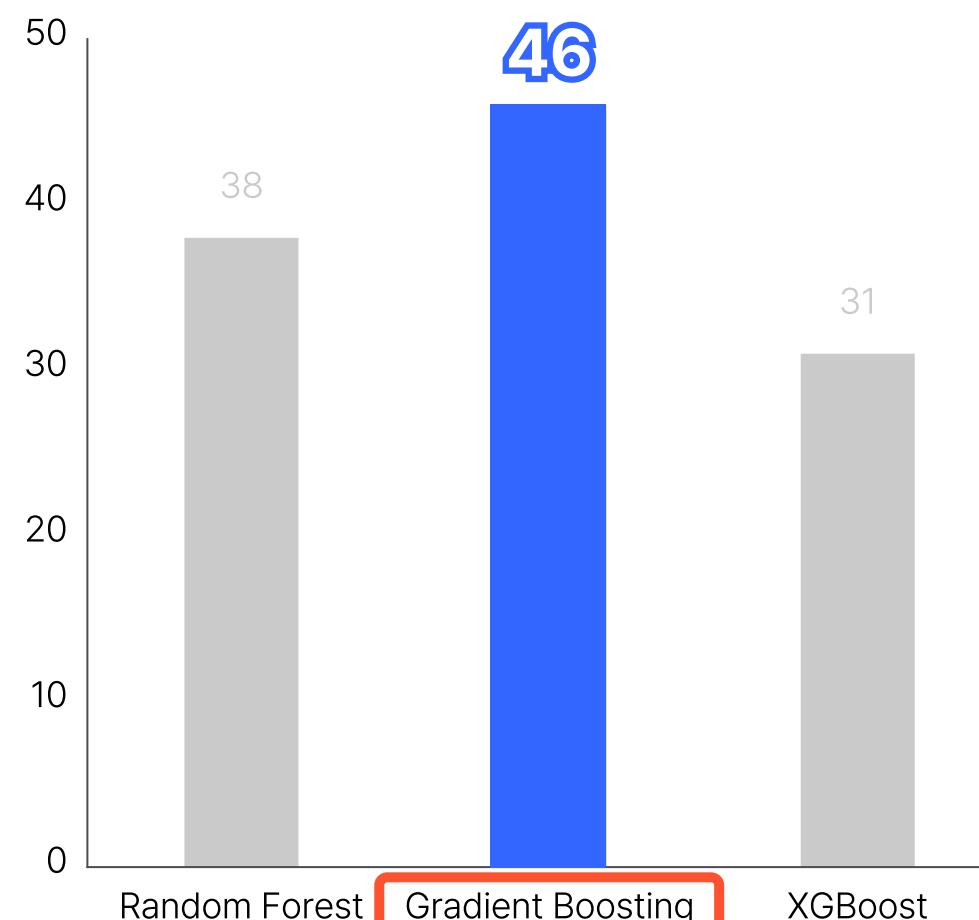
R <sup>2</sup> (모델 성능 평가 지표)	
Random Forest	0.378069
Gradient Boosting	0.458389
XGBoost	0.305196

AirPods Pro2

R <sup>2</sup>	
Random Forest	0.243323
Gradient Boosting	0.086759
XGBoost	0.010367

Iphone 15

R <sup>2</sup>	
Random Forest	0.086421
Gradient Boosting	0.235246
XGBoost	0.272450



중고 가격은 다양한 요인들이 복합적으로 작용하여  
특정한 한 가지 모델만으로 모든 상황에 우수한 성능을 기대하기 어려움

→ 최고의 성능을 나타내는 모델을 사용하기로 결정

# 다중 모델 사용의 이점

1

## 최적 성능 유지 및 모델의 상호 보완 가능

- 모델 간의 상호 보완적 효과
- 특정 모델이 제품마다 성능이 상이
- 여러 모델 사용 시 각 상황의 최적 모델을 사용하여 전반적으로 더 나은 성능 유지 가능

2

## 데이터 특성에 따른 최적화된 대응 가능

- 중고 거래 데이터는 상품의 종류, 상태, 가격 등 다양한 요인에 따라 데이터 특성 다름
- 하나의 모델로 모든 데이터의 복잡성 포착 어려움

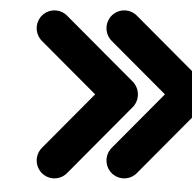
3

## 적절한 규모의 데이터셋으로 인한 다중 모델 사용의 효율성

- 데이터 크기(카테고리별 최대 5000개 이하)가 비교적 크지 않음
- 다중 모델 사용 시 학습 시간과 계산 비용 부담 낮음

# 상품 상태별 예측 가격 범위 설정 이유

상품의 상태를 모델에 입력 시  
단일 예측 가격 출력



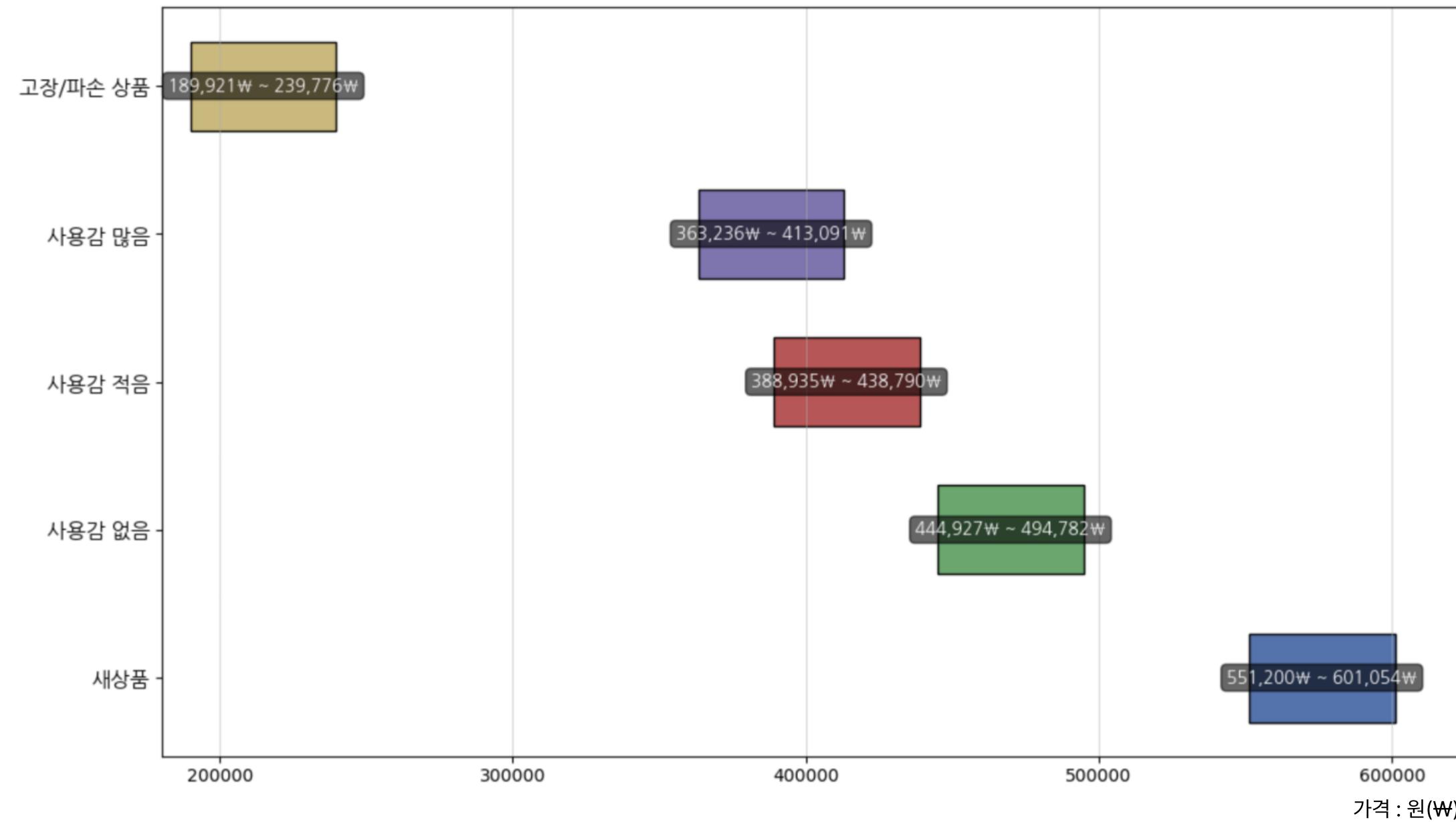
중고 거래는 상품의 상태, 종류, 가격 등 다양한 변수 작용  
단일 예측 가격보다 가격 범위가 더 신뢰성 있는 정보를 제공

예측 오차를 반영한 가격 범위를 제시하여, 사용자에게 현실적인 가격 범위 제공

# 상품 상태별 예측 가격 범위 설정

**MAE**  
(평균 절대 오차)

예측 모델이 실제 값과 얼마나 차이 나는지 평균적으로 보여주는 지표



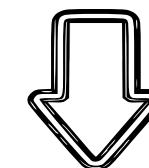
MAE를 활용하여 예측 가격의 오차 측정  
예측 가격에  $\pm$ MAE를 적용해 신뢰구간 설정

# 가격 범위 생성 과정

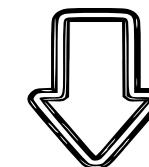
해당 카테고리의 중고 물품 데이터 가져옴



3가지 모델을 학습한 후,  $R^2$  값을 비교하여 가장 성능이 좋은 모델을 선택

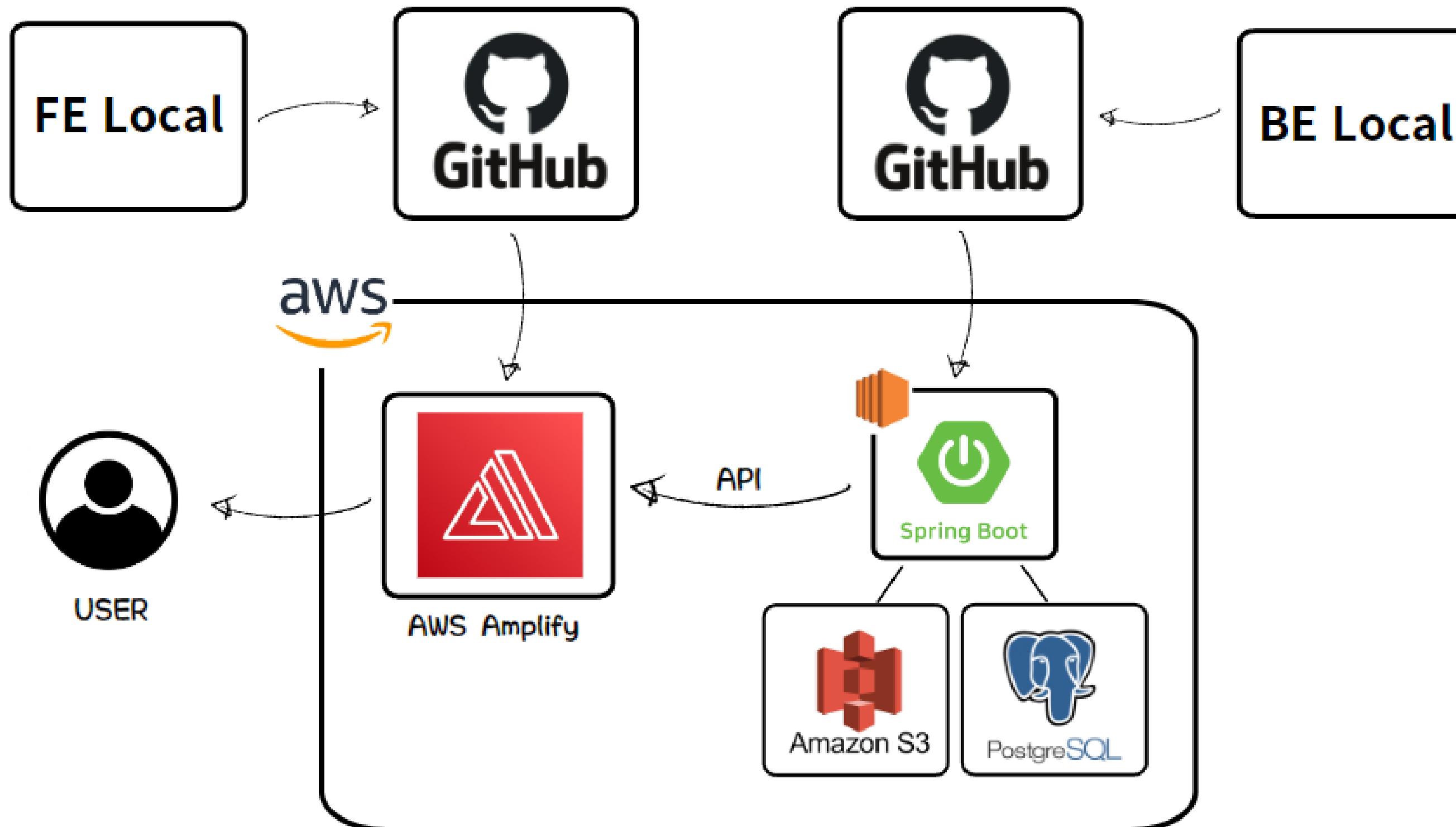


상태별로 평균 절대오차(MAE)를 활용해 예측된 가격에 기반한 가격 범위 생성



생성된 가격 범위 저장하여 사용자에게 제공

# 가격 예측 모델을 이용한 웹서비스



# 가격 예측 모델을 이용한 웹서비스

특정 카테고리의 아이템 수가 20개 이상 추가될 때마다 자동으로 가격 예측 모델이 업데이트 되는 비동기 처리 시스템을 구축



사용자는 항상 최신 가격 정보 확인 가능

- ① 비동기 처리 방식 덕분에 데이터 처리 속도가 빨라져 전체 시스템이 더욱 효율적
- ② 빠르고 정확한 가격 예측 결과를 제공하여 사용자 만족도 증대

# 가격 예측 모델을 이용한 웹서비스

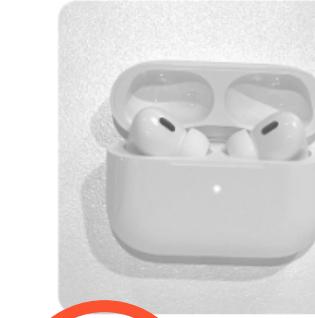
실시간 추천 아이템 !



✓ 에어팟 맥스 팝니다 (상태좋음)  
580,000원  
2024-11-12



✓ 아이폰 15프로 256GB  
1,200,000원  
2024-11-12



✓ 애플 에어팟 프로 2  
230,000원  
2024-11-12

추천 가격 범위내에 있을 때



✓ 에어팟 프로2 본체  
60,000원  
2024-11-12



✓ Apple 2024 에어팟 맥스 노이즈 캔슬링  
블루투스 헤드폰  
600,000원  
2024-11-12



✓ 에어팟 프로2 왼쪽  
50,000원  
2024-11-12

방금 등록된 상품



에어팟 맥스  
350,000원  
2024-11-12



아이폰 15 화이트 팝니다  
600,000원  
2024-11-12



✓ 에어팟 프로2 왼쪽  
50,000원  
2024-11-12



케이스티파이 아이폰 15 프로  
32,000원  
2024-11-12



✓ 에어팟 프로2 본체  
60,000원  
2024-11-12

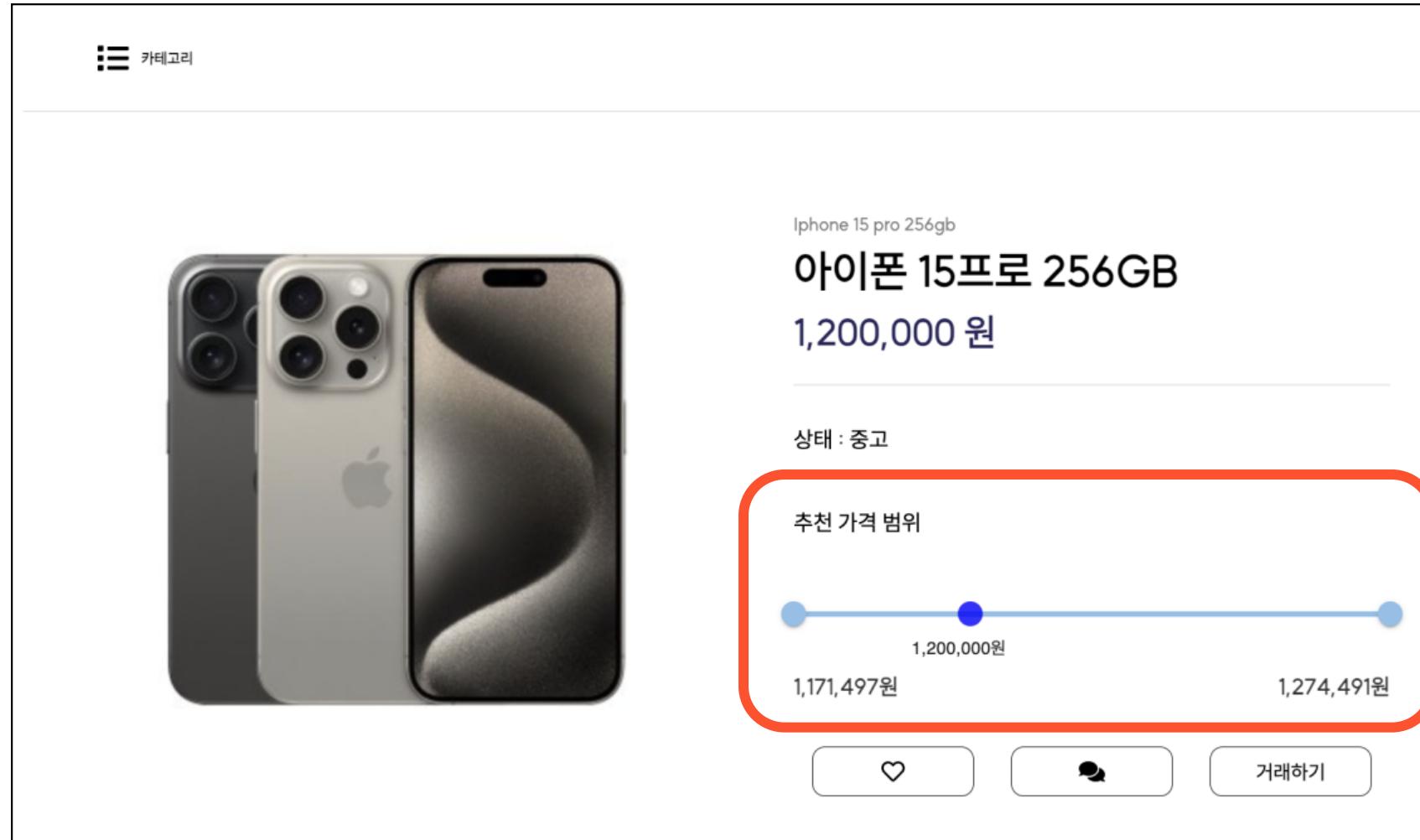


✓ 아이폰 15프로 256GB  
1,200,000원  
2024-11-12

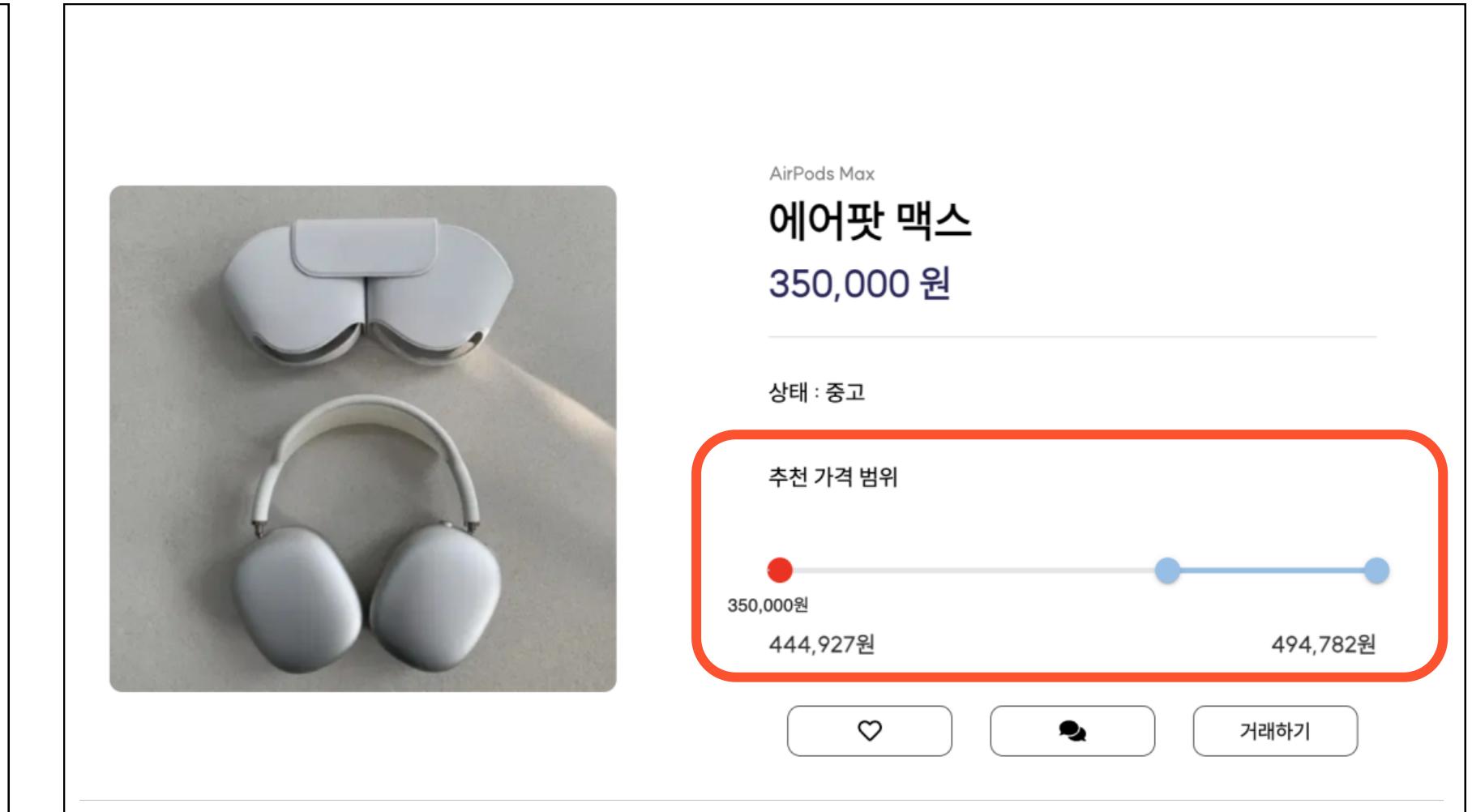
사용자들에게 해당 물품이 현재 시장에서 적정한 가격에 판매되고 있음을 시각적으로 나타냄

# 가격 예측 모델을 이용한 웹서비스

추천 가격 범위 내에 있을 때



추천 가격 범위 외에 있을 때



제품의 추천 가격 범위를 한눈에 확인할 수 있음

# 가격 예측 모델을 이용한 웹서비스

판매자가 카테고리와 상품 상태를 선택하여, 조건에 맞는 적정 가격을 쉽게 찾을 수 있는 필터 기능 제공

카테고리

웨어러블 ▾ 에어팟 ▾ 에어팟 프로2 ▾

상품상태

- 새 상품
- 사용감 적음
- 사용감 많음
- 중고
- 고장/파손

230,000원

258,325원

202,145원

카테고리

웨어러블 ▾ 에어팟 ▾ 에어팟 프로2 ▾

상품상태

- 새 상품
- 사용감 적음
- 사용감 많음
- 중고
- 고장/파손

280,000원

258,325원

202,145원

가격

230000

가격

280000

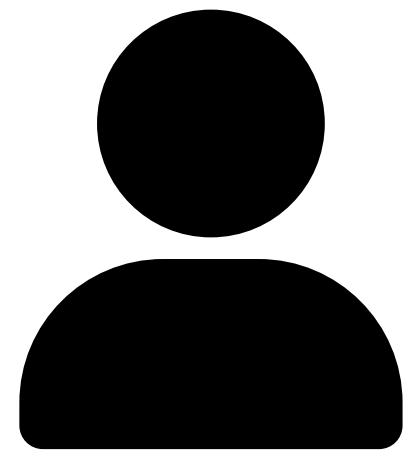
가격 추천 범위 밖입니다.

판매자는 상품의 상태에 맞는 적정한 가격대를 쉽게 파악 가능

# 기대효과

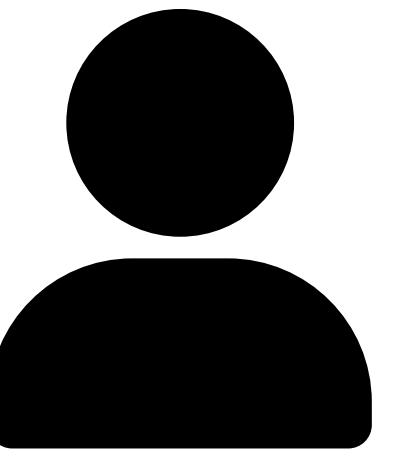
- 1** 추천 가격을 제시함으로써, 판매자가 적정 가격을 쉽게 파악하고 설정할 수 있어 가격 결정에 대한 부담이 줄어듦
- 2** 구매자와 판매자 모두 합리적인 가격을 기준으로 거래를 진행할 수 있어 거래가 더 신속하게 성사될 가능성이 높아짐
- 3** 모든 사용자가 참고할 수 있는 기준 가격이 제시됨에 따라, 비정상적으로 높은 가격이나 낮은 가격의 거래가 줄어들어 중고 거래 시장의 가격이 안정화됨

# 팀원 소개



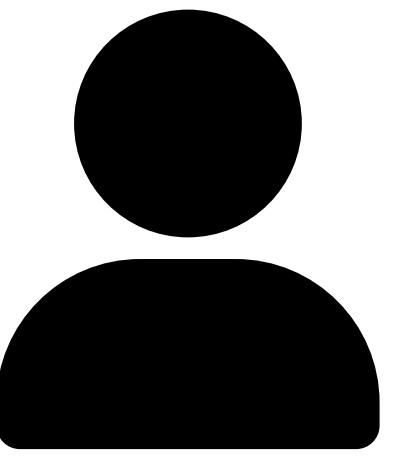
박지원

**Project Manager,  
Data Science**



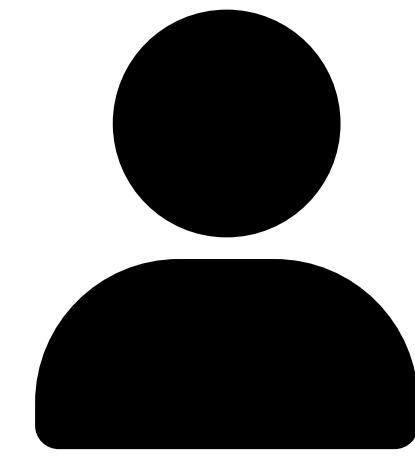
김민지

**Front-end**



배성환

**Back-end**



이하준

**Back-end**

**감사합니다**

# **Q & A**