

---

# 재무제표 기반 5대 게임사 주가 요인 분석/예측

202144098 이상현



# DATA ANALYTICS

## WHY?

**국민적 관심 증대** : 코로나19 팬데믹 이후 투자 시장 참여자의 급증과 성공적인 투자를 향한 높은 열망.

**분석 대상 선정 배경** : 평소 높은 관심도를 가진 산업이자 대한민국을 대표하는 5대 게임 기업을 분석 대상으로 선정.

**산업적 가치** : 국민 취미 생활을 넘어선 높은 성장 잠재력과 지속 가능한 가치에 대한 진단 필요.

---

# WHAT?

**분석 대상 산업:** 게임 산업의 높은 민감도

(국가 규제, 국제 정세 등 비재무적 요인에 주가가 민감).

**문제 정의:** 기업의 재무 건전성과 외부 이슈가 주가 변동에 미치는 복합적인 영향을 규명해야 할 필요성.

**분석 범위:** 5대 게임사의 재무 현황, 주가 추이, 그리고 굵직한 사건 간의 연관성 규명.

---

# HOW?

**핵심 방법론:** 가장 통상적이고 객관적인 재무제표 분석을 핵심 방법론으로 채택.

**분석 과정:** 재무 현황 및 주가 추이를 시각화하고 사건과의 상관관계를 분석하여 과거 변동 요인을 진단.

**최종 목표:** 재무제표 등 정량 데이터만을 활용한 머신러닝(ML) 기반의 주가 예측 모델 구축.

---

# WHEN?

데이터 수집 기간: 2018년 1분기 ~ 2025년 3분기까지의 데이터 확보

---

# WHERE?

펀더멘털 지표: 금융감독원 전자공시시스템 (DART) API.

시장 가치 지표: PyKRX 라이브러리 (KRX: 한국거래소)

# 데이터 수집 정기보고서

요청키	명칭	타입	필수여부	값설명
crtfc_key	API 인증키	STRING(40)	Y	발급받은 인증키(40자리)
corp_code	고유번호	STRING(8)	Y	공시대상회사의 고유번호(8자리) ※ 개발가이드 > 공시정보 > 고유번호 참고
bsns_year	사업연도	STRING(4)	Y	사업연도(4자리) ※ 2015년 이후 부터 정보제공
reprt_code	보고서 코드	STRING(5)	Y	1분기보고서 : 11013 반기보고서 : 11012 3분기보고서 : 11014 사업보고서 : 11011

- 쿼리 실행을 위해서 고유번호를 얻어야 함.

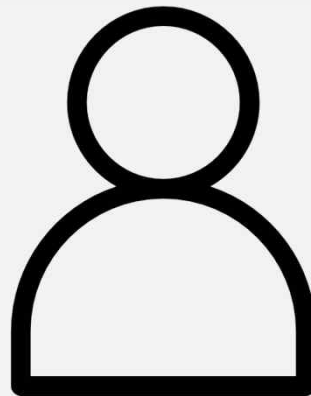
# 데이터 수집 고유번호



쿼리



[https://opendart.fss.or.kr/api/corpCode.xml?certfc\\_key](https://opendart.fss.or.kr/api/corpCode.xml?certfc_key)





# Result

결측치 발생!

	corp_code	corp_name	stock_code	modify_date
0	00441854	넷마블		20170630
1	00760971	크래프톤	259980	20231027
2	01096341	넥슨게임즈	225570	20240110
3	01137383	카카오게임즈	293490	20240402
4	00261443	엔씨소프트	036570	20240329

```
import requests
import zipfile
import io
import xml.etree.ElementTree as ET
import pandas as pd

# API 인증키 입력
crtfc_key = 
url = f'https://opendart.fss.or.kr/api/corpCode.xml?crtfc_key={crtfc_key}'

# API 요청 및 ZIP 파일 받기
response = requests.get(url)
zip_bytes = io.BytesIO(response.content)

# ZIP 압축 해제
with zipfile.ZipFile(zip_bytes) as z:
    xml_filename = z.namelist()[0]
    with z.open(xml_filename) as xml_file:
        xml_content = xml_file.read()

# XML 파싱
root = ET.fromstring(xml_content)

# 기업코드 리스트 추출
corp_list = []
empty_list = []
for corp in root.findall('list'):
    c_name = corp.findtext('corp_name')
    if c_name=='넥슨게임즈' or c_name=='엔씨소프트' or c_name=='크래프톤' or c_name=='카카오게임즈' or c_name=='넷마블':
        corp_list.append({
            'corp_code': corp.findtext('corp_code'),
            'corp_name': c_name,
            'stock_code': corp.findtext('stock_code'),
            'modify_date': corp.findtext('modify_date')
        })

# DataFrame으로 변환
corp_target_df = pd.DataFrame(corp_list)
print(corp_target_df.head())
```

# 데이터 수집

## 공시검색

요청키	명칭	타입	필수여부	값설명
crtfc_key	API 인증키	STRING(40)	Y	발급받은 인증키(40자리)
corp_code	고유번호	STRING(8)	N	공시대상회사의 고유번호(8자리) ※ <a href="#">개발가이드 &gt; 공시정보 &gt; 고유번호 참고</a>
bgn_de	시작일	STRING(8)	N	검색시작 접수일자(YYYYMMDD) 1) 기본값 : 종료일(end_de) 2) 고유번호(corp_code)가 없는 경우 검색기간은 3개월로 제한
end_de	종료일	STRING(8)	N	검색종료 접수일자(YYYYMMDD) 1) 기본값 : 당일
pblntf_ty	공시유형	STRING(1)	N	A : 정기공시 B : 주요사항보고 C : 발행공시 D : 지분공시 E : 기타공시 F : 외부감사관련 G : 펀드공시 H : 자산유동화 I : 거래소공시 J : 공정위공시
pblntf_detail_ty	공시상세유형	STRING(4)	N	(※ 상세 유형 참조 : pblntf_detail_ty)
page_no	페이지 번호	STRING(5)	N	페이지 번호(1~n) 기본값 : 1
page_count	페이지 별 건수	STRING(3)	N	페이지당 건수(1~100) 기본값 : 10, 최대값 : 100

```

import requests

finance_filepath = None
stock_filepath = None

page_count = max

webpages = []

bsns_year = ['2018', '2019', '2020', '2021', '2022', '2023', '2024', '2025']

# 기업별 공시 검색 및 데이터 수집
for idx in corp_target_df.index:
    corp_code = corp_target_df.loc[idx, 'corp_code']
    corp_name = corp_target_df.loc[idx, 'corp_name']

# 연도별 검색 기간 설정 및 API URL 생성
for year in bsns_year:
    bgn_de = year + '0501' # 1분기 보고서 발행월
    end_de = str(int(year) + 1) + '0430' # 다음 해 3월까지

    url = [
        f'https://opendart.fss.or.kr/api/list.xml?crtfc_key={crtfc_key}',
        f'&corp_code={corp_code}&pbIntf_ty=A',
        f'&bgn_de={bgn_de}&end_de={end_de}',
        f'&page_count=100',
        f'&pbIntf_detail_ty=A001&pbIntf_detail_ty=A002&pbIntf_detail_ty=A003'
    ]

# API 요청 및 응답 디코딩
resp = requests.get(url)
content = resp.content.decode('UTF-8')

# 수집 데이터 리스트에 추가
webpages.append({
    'corp_name': corp_name,
    'year': year,
    'content': content
})

# 수집 결과 확인
for page in webpages:
    print(f"{page['corp_name']} - {page['year']}")
    print(page['content'][:500]) # 앞부분만 출력

```

# Result

넥슨게임즈 - 2018 <?xml version="1.0" encoding="UTF-8" standalone="yes"?><result><status>000</status><message>정상</message><page\_no>1</page\_no><page\_count>100</page\_count><total\_count>5</total\_count><total\_page>1</total\_page><list><corp\_code>01096341</corp\_code><corp\_name>넥슨게임즈</corp\_name><stock\_code>225570</stock\_code><corp\_cls>K</corp\_cls><report\_nm>사업보고서(2018.12)</report\_nm><rcept\_no>20190401003993</rcept\_no><flr\_nm>넥슨게임즈</flr\_nm><rcept\_dt>20190401</rcept\_dt><rm></rm></list><list><corp\_code>01096341</corp\_code>

# 데이터 수집 XML 파싱

응답키	명칭	출력설명
list		
corp_cls	법인가분	법인가분 : Y(유가), K(코스닥), N(코넥스), E(기타)
corp_name	종목명(법인명)	공시대상회사의 종목명(상장사) 또는 법인명(기타법인)
corp_code	고유번호	공시대상회사의 고유번호(8자리)
stock_code	종목코드	상장회사의 종목코드(6자리)
report_nm	보고서명	공시구분+보고서명+기타정보 [기타정보] : 본 보고서명으로 이미 제출된 보고서의 기재내용이 변경되어 제출된 것임 [첨가] : 본 보고서명으로 이미 제출된 보고서의 첨부내용이 변경되어 제출된 것임 [첨가] : 본 보고서명으로 이미 제출된 보고서의 첨부서류가 추가되어 제출된 것임 [변경] : 본 보고서명으로 이미 제출된 보고서의 유동화계획이 변경되어 제출된 것임 [연장] : 본 보고서명으로 이미 제출된 보고서의 신탁계약이 연장되어 제출된 것임 [발행조건] : 본 보고서명으로 이미 제출된 보고서의 유가증권 발행조건이 확정되어 제출된 것임 [정명령부과] : 본 보고서에 대하여 금융감독원이 정정명령을 부과한 것임 [제출요구] : 본 보고서에 대하여 금융감독원이 정정제출요구를 부과한 것임
rcpt_no	접수번호	접수번호(14자리) ※ 공시뷰어 연결에 이용예시 - PC용 : <a href="https://dart.fss.or.kr/dsaf001/main.do?rcpNo=접수번호">https://dart.fss.or.kr/dsaf001/main.do?rcpNo=접수번호</a>
flr_nm	공시 제출인명	공시 제출인명
rcpt_dt	접수일자	(MDD)
rm	비고	공 : 본 공시사항은 공정거래위원회 소관임 연 : 본 보고서는 연결부분을 포함한 것임 정 : 본 보고서 제출 후 정정신고가 있으니 관련 보고서를 참조하시기 바람 철회 : 본 보고서는 철회(간주)되었으니 관련 철회신고서(철회간주안내)를 참고하시기 바람

결측치 해결

	corp_code	corp_name	stock_code	modify_date
0	00441854	넷마블		20170630
1	00760971	크래프톤	259960	20231027
2	01096341	넥슨게임즈	225570	20240110
3	01137383	카카오게임즈	293490	20240402
4	00261443	엔씨소프트	036570	20240329

# 데이터 수집 정기보고서

	corp_name	corp_code	bsns_year	reprt_code	account_nm	fs_div	sj_div	thstrm_nm	thstrm_dt	thstrm_ament	frmtrm_nm	frmtrm_dt	frmtrm_ament	currency	rcept_no	rcept_dt	thstrm_ament_int
0	넥슨게임즈	01096341	2018	11011	유동자산	OFS	BS	제 4 기	2018.12.31 현재	22,159,210,178	제 3 기	2017.12.31 현재	35,726,056,445	KRW	20190401003993	20190401	22159210178
1	넥슨게임즈	01096341	2018	11011	비유동자산	OFS	BS	제 4 기	2018.12.31 현재	7,256,714,003	제 3 기	2017.12.31 현재	7,262,731,234	KRW	20190401003993	20190401	7256714003
2	넥슨게임즈	01096341	2018	11011	자산총계	OFS	BS	제 4 기	2018.12.31 현재	29,415,924,181	제 3 기	2017.12.31 현재	42,988,787,679	KRW	20190401003993	20190401	29415924181
3	넥슨게임즈	01096341	2018	11011	유동부채	OFS	BS	제 4 기	2018.12.31 현재	4,510,336,944	제 3 기	2017.12.31 현재	5,313,282,771	KRW	20190401003993	20190401	4510336944
4	넥슨게임즈	01096341	2018	11011	비유동부채	OFS	BS	제 4 기	2018.12.31 현재	6,237,056,529	제 3 기	2017.12.31 현재	8,167,999,100	KRW	20190401003993	20190401	6237056529
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3767	크래프톤	00760971	2025	11014	당기순이익 (손실)	CFS	IS	제 19 기3분기	2025.01.01 ~ 2025.09.30	369,449,615,687	제 18 기3분기	2024.01.01 ~ 2024.09.30	121,360,898,732	KRW	20251114002295	20251114	369449615687
3768	크래프톤	00760971	2025	11014	총포괄손익	CFS	IS	제 19 기3분기	2025.01.01 ~ 2025.09.30	397,064,976,492	제 18 기3분기	2024.01.01 ~ 2024.09.30	74,174,180,148	KRW	20251114002295	20251114	397064976492
3769	크래프톤	00760971	2025	11014	영업비용	CFS	IS	제 19 기3분기	2025.01.01 ~ 2025.09.30	522,015,105,876	제 18 기3분기	2024.01.01 ~ 2024.09.30	394,854,576,984	KRW	20251114002295	20251114	522015105876

1784 rows × 17 columns

---

# 데이터 수집 정기보고서 전처리

**데이터 클리닝**: 금액 데이터를 숫자형으로 변환하고, 데이터의 공백과 중복을 제거하며, 결측된 분기 실적을 후행 연도의 데이터로 보간.

**핵심 가공**: 연간 사업보고서(4분기 누적)에서 1~3분기 실적을 차감하여  
4분기만의 순수한 실적을 정확하게 산출

**저장**: 회사명, 연도-분기, 계정명을 조합하여 분석에 용이한 인덱스를 설정하고,  
이 정제된 데이터를 데이터 프레임으로 만든 뒤 CSV 파일로 저장.

---

# 데이터 수집 정기보고서

	corp_name	bsns_year	reprt_code	account_nm	thstrm_amount	frmtrm_amount	currency	rcept_dt	year_quarter
corp_yqan									
넥슨게임즈-2018-Q1-당기순이익(손실)	넥슨게임즈	2018	11013	당기순이익(손실)	4.210047e+08	2.957616e+09	KRW	20180515	2018-Q1
넥슨게임즈-2018-Q1-매출액	넥슨게임즈	2018	11013	매출액	8.590205e+09	6.856382e+09	KRW	20180515	2018-Q1
넥슨게임즈-2018-Q1-법인세차감전순이익	넥슨게임즈	2018	11013	법인세차감전순이익	4.210047e+08	2.343760e+09	KRW	20180515	2018-Q1
넥슨게임즈-2018-Q1-부채총계	넥슨게임즈	2018	11013	부채총계	1.316007e+10	1.348128e+10	KRW	20180515	2018-Q1
넥슨게임즈-2018-Q1-영업이익	넥슨게임즈	2018	11013	영업이익	2.987293e+08	2.366868e+09	KRW	20180515	2018-Q1
...	...	...	...	...	...	...	...	...	...
크래프톤-2025-Q3-자본금	크래프톤	2025	11014	자본금	4.927645e+09	4.925891e+09	KRW	20251114	2025-Q3
크래프톤-2025-Q3-자본총계	크래프톤	2025	11014	자본총계	7.279704e+12	6.829132e+12	KRW	20251114	2025-Q3
크래프톤-2025-Q3-자산총계	크래프톤	2025	11014	자산총계	8.441269e+12	7.919464e+12	KRW	20251114	2025-Q3

1168 rows × 9 columns

# 데이터 수집 주가 데이터

corp_code	corp_name	stock_code	modify_date
00760971	크래프톤	259960	20231027
01096341	넥슨게임즈	225570	20240110
01137383	카카오게임즈	293490	20240402
00261443	엔씨소프트	036570	20240329
00904672	넷마블	251270	20250917

```
from pykrx import stock
import pandas as pd

all_stock_data = []

for stock_code in corp_codes:
    # 회사명 가져오기
    corp_name = corp_target_df.loc[corp_target_df['stock_code'] == stock_code, 'corp_name'].iloc[0]
    print(stock_code, '-', corp_name)

    # 2018년부터 2025년 3분기까지의 주가 데이터 가져오기
    df = stock.get_market_ohlc("20180101", "20250930", stock_code)

    # 결측치 제거
    df.dropna(how='any', inplace=True)

    # 회사명과 종목코드 추가
    df['corp_name'] = corp_name
    df['stock_code'] = stock_code
    df['날짜'] = df.index

    # 전체 통합용 리스트에 추가
    all_stock_data.append(df)

# 모든 회사 데이터를 하나로 합치기
final_stock_df = pd.concat(all_stock_data).reset_index(drop=True)

# 전체 통합 CSV 저장
final_stock_filepath = '/content/AllCompanies_Stock.csv'
final_stock_df.to_csv(final_stock_filepath, encoding='utf-8-sig', index=False)

print(final_stock_df.head())
```



# 데이터 수집

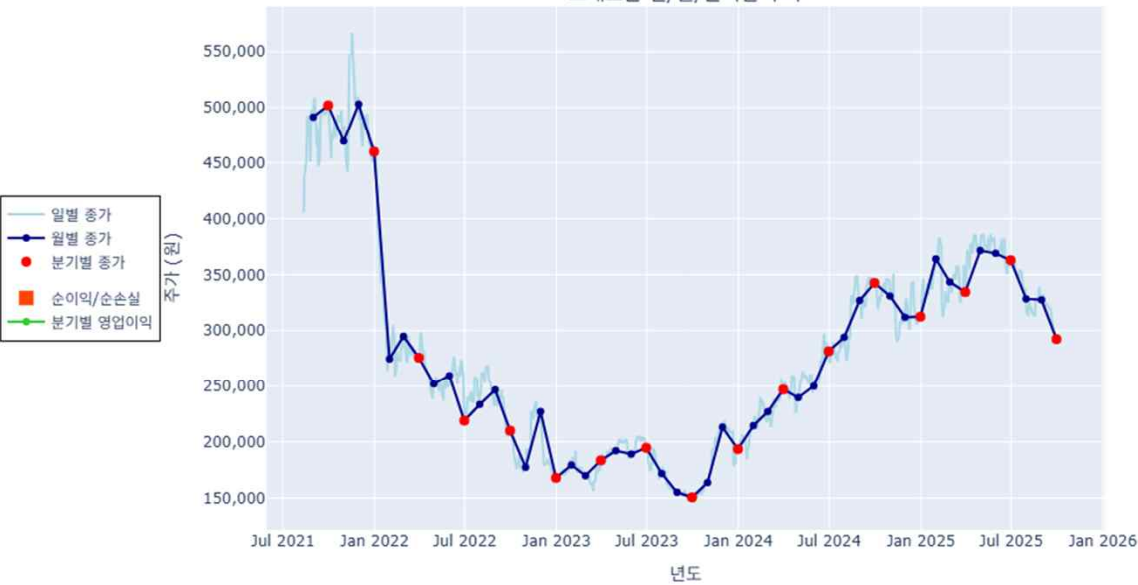
## 주가 데이터

	시가	고가	저가	종가	거래량	등락률	corp_name	stock_code	날짜
0	444500	446000	405000	407000	1647759	-10.352423	크래프톤	259960	2021-08-11
1	414000	420500	402000	406000	947958	-0.245700	크래프톤	259960	2021-08-12
2	415000	445500	408500	437000	1669847	7.635468	크래프톤	259960	2021-08-13
3	433000	460000	423000	451500	1189500	3.318078	크래프톤	259960	2021-08-17
4	449000	480000	437000	469000	1168167	3.875969	크래프톤	259960	2021-08-18
...	...	...	...	...	...	...	...	...	...
7963	64200	64500	62000	62200	130981	-2.660407	넷마블	251270	2025-09-26
7964	63300	64300	62400	63000	99704	1.286174	넷마블	251270	2025-09-29
7965	62800	63500	61700	61700	84570	-2.063492	넷마블	251270	2025-09-30

7966 rows × 9 columns

크래프톤 주가 및 재무제표 그래프

크래프톤 일/월/분기별 주가

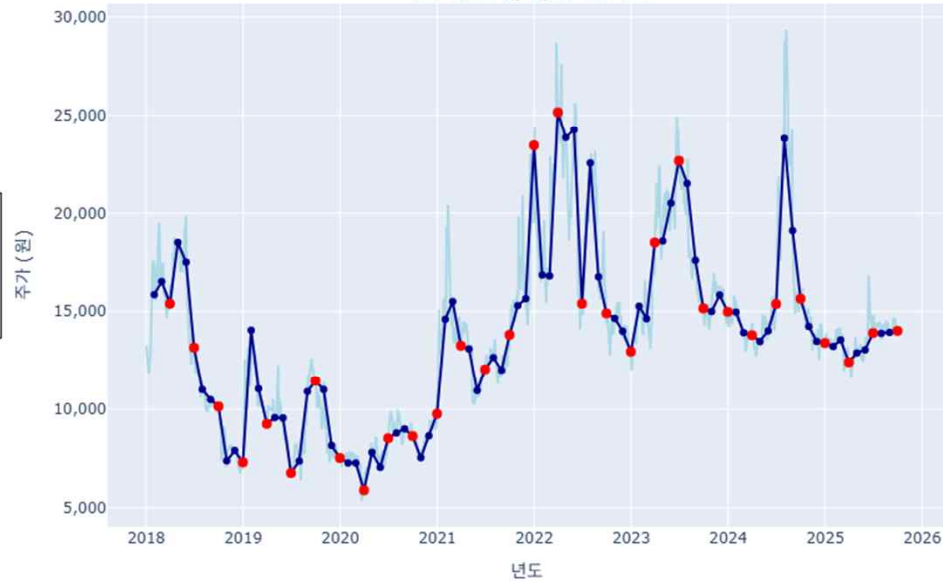


크래프톤 분기별 매출현황

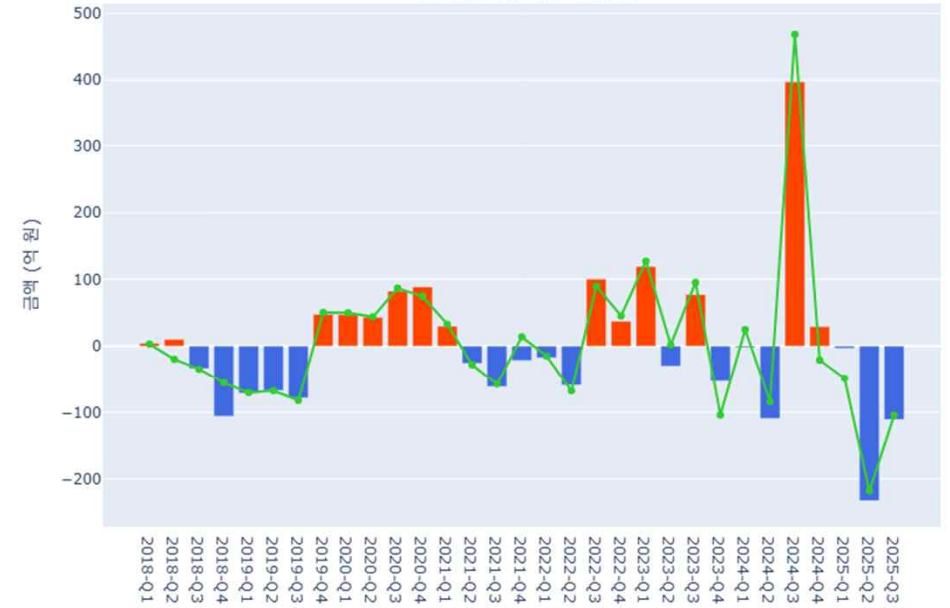


넥슨게임즈 주가 및 재무제표 그래프

넥슨게임즈 일/월/분기별 주가

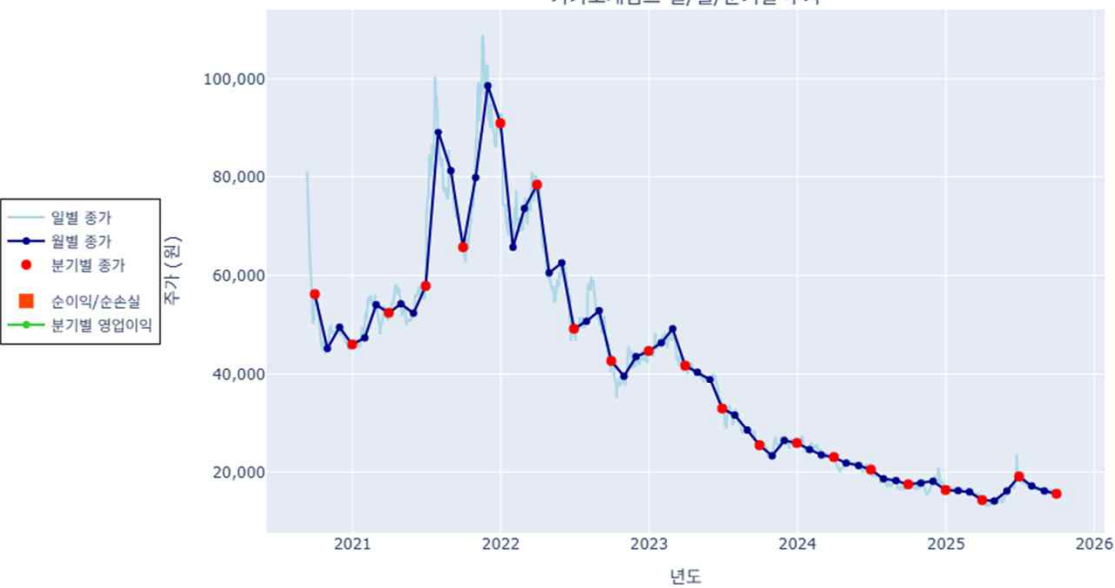


넥슨게임즈 분기별 매출현황



카카오게임즈 주가 및 재무제표 그래프

카카오게임즈 일/월/분기별 주가

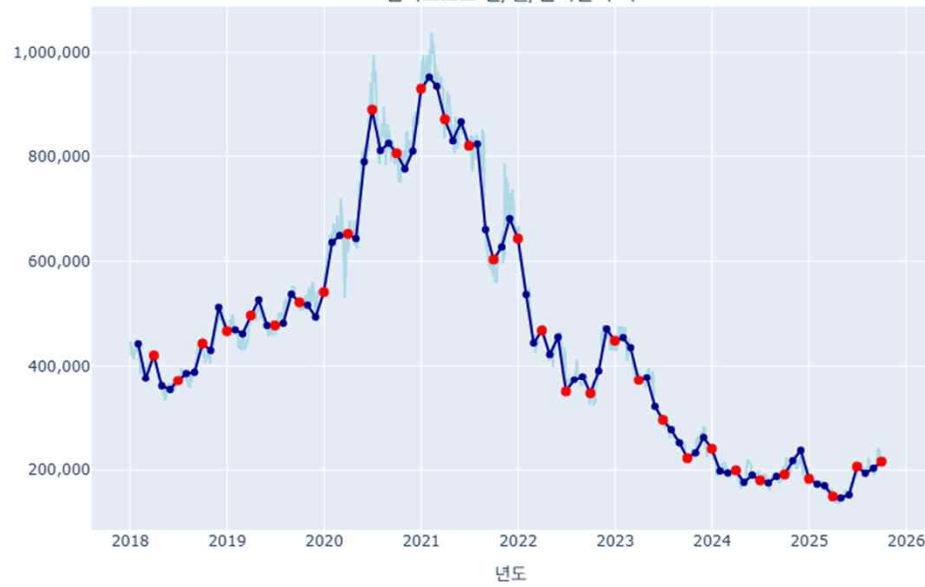


카카오게임즈 분기별 매출현황

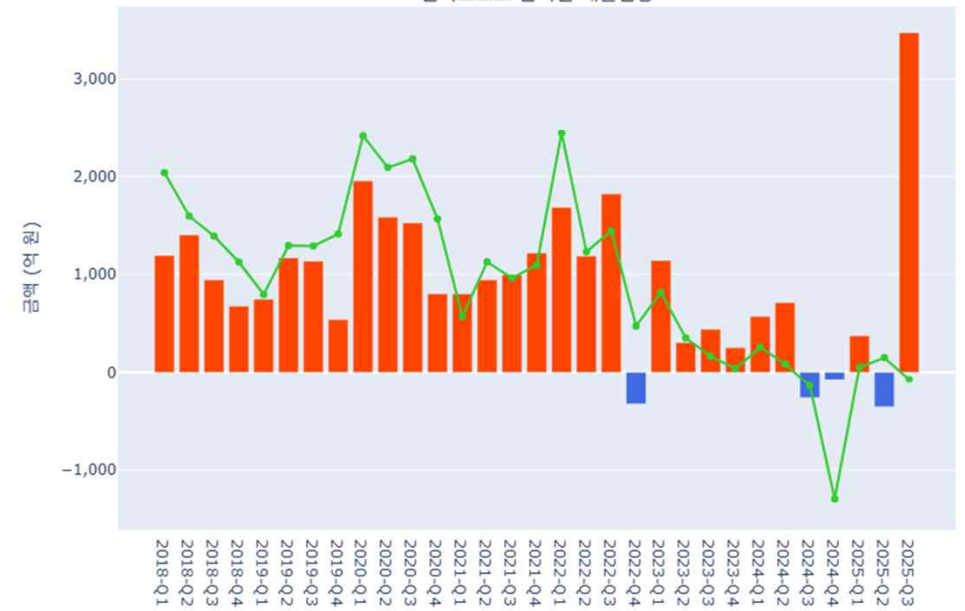


엔씨소프트 주가 및 재무제표 그래프

엔씨소프트 일/월/분기별 주가



엔씨소프트 분기별 매출현황

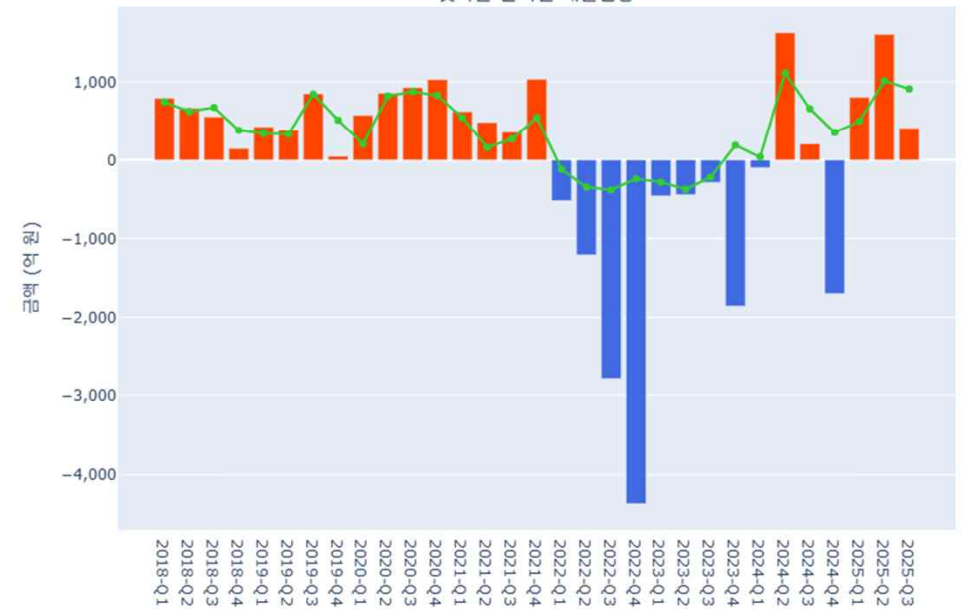


넷마블 주가 및 재무제표 그래프

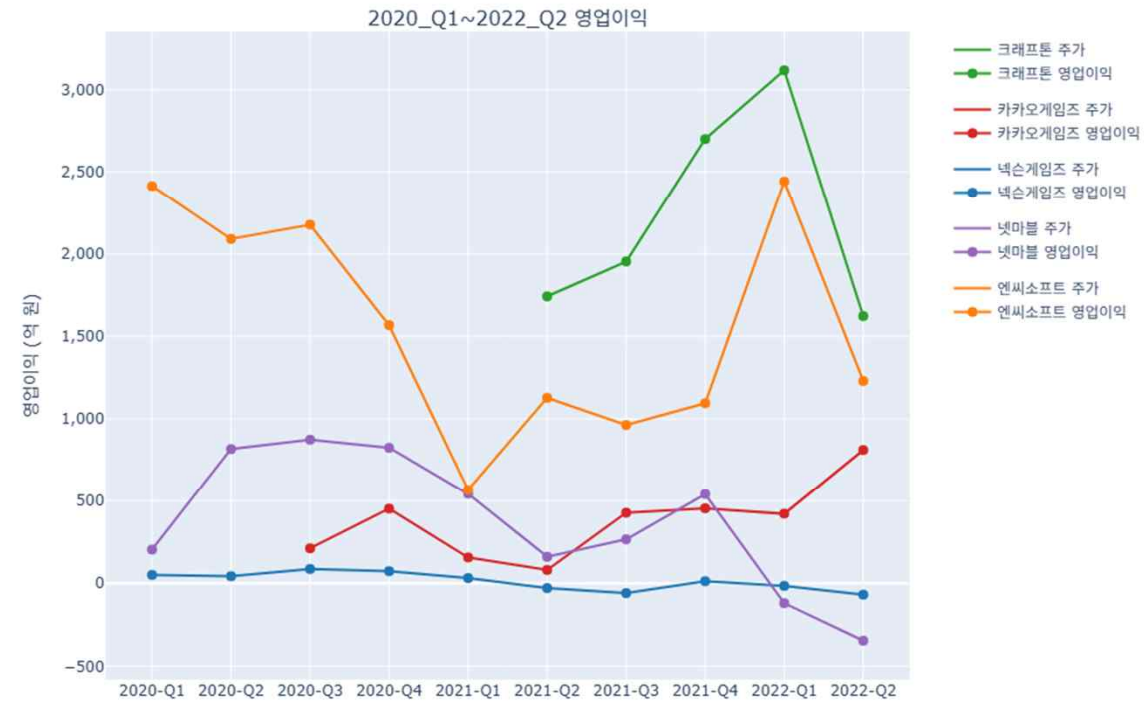
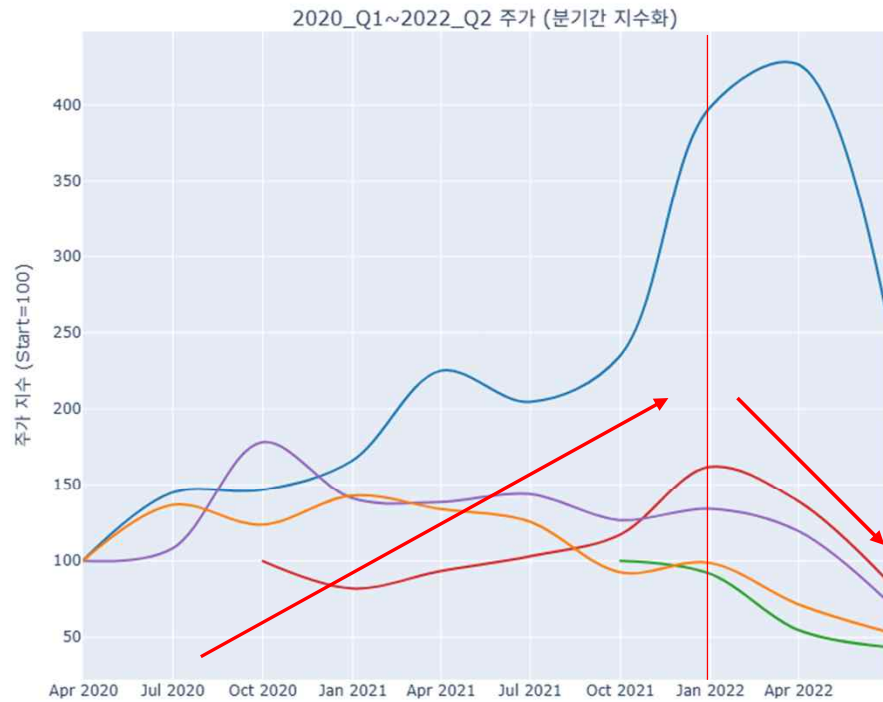
넷마블 일/월/분기별 주가



넷마블 분기별 매출현황



## 2020\_Q1~2022\_Q2 기업 비교 분석



# WHY?

**대규모 유입:** 팬데믹 기간 중 재택근무 확산은 게임 이용자의 대규모 유입을 유도함.

**시장 재평가:** 이러한 구조적 변화는 해당 기업들의 미래 실적 기대치를 끌어올리는 결정적 요인으로 작용하여 시장 재평가를 이끌어냄.

## 코로나 발생 이후 재택근무 비율 추이

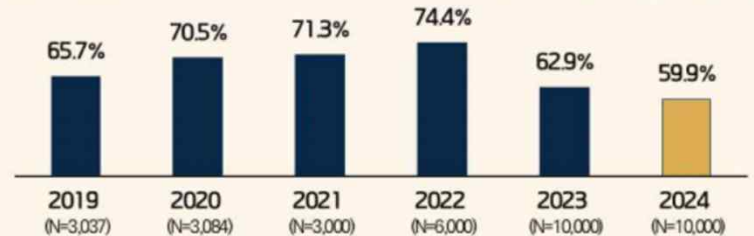
Shiftree

\* 2020~2021년 재택근무 데이터 기준



## 게임 이용률

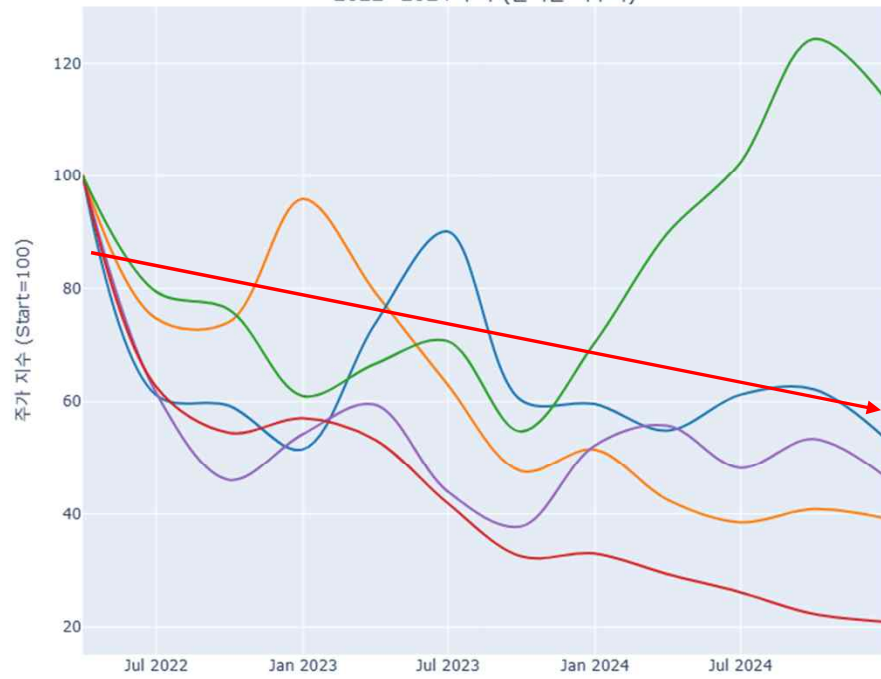
[BASE: 2019년~2023년 10~64세, 2024년 10~69세]



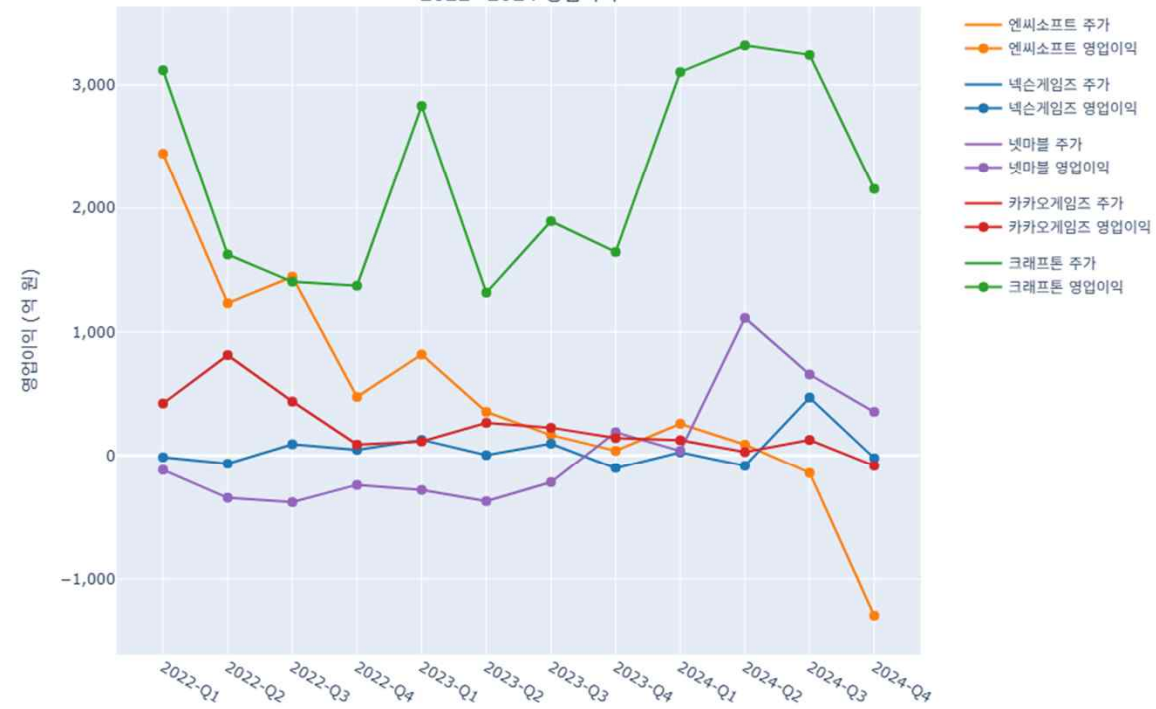


## 2022~2024 기업 비교 분석

2022~2024 주가 (분기간 지수화)



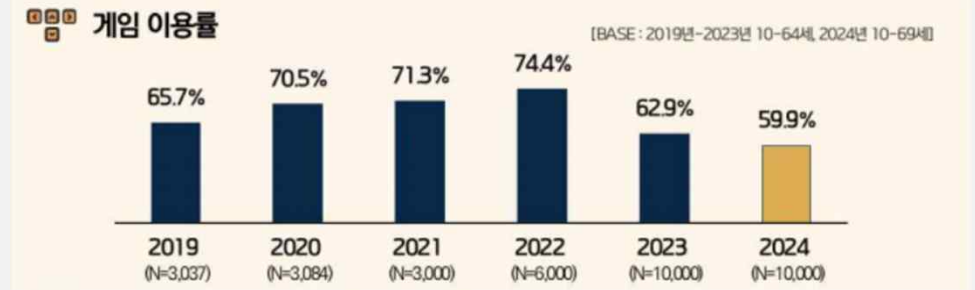
2022~2024 영업이익



# WHY?

**특수 소멸**: 재택 근무 종료로 인한  
유입 및 기존 유저 이탈

**물가 상승**: 고물가로 인한 필수재  
지출 우선(게임 소비 축소)



---

# CONCLUSION

**상승과 하락:** 한국 게임 산업은 2020년부터 2024년까지 거시경제 환경 변화에 따라 황금기와 냉각기를 경험.

**낮은 변동성:** 오락 및 문화 지수는 2018년부터 2024년까지 변동폭이 타 지수에 비해 낮은 편.

**거시경제 민감성:** 2020년~2022년 상반기의 주가 지수 상승 폭은 영업 이익의 실제 성장 폭을 압도하며 선행함. 반면, 2022년 이후 금리 인상 및 물가 급등기에는 주가 하락 폭이 영업 이익 감소 폭보다 컸음.

---

# MACHINE LEARNING

## HOW?

2025년 10월 ~ 11월 주가 예측

### 1. 데이터 수집 및 가공

- 주가 데이터 : 2018년부터 2025년 3분기까지의 일별 주가 (시가, 고가, 저가, 종가, 거래량)를 수집.
- 재무 데이터 : 과거 분기별 실적(매출액, 영업이익, 순이익, 부채)을 수집.
- 지표 생성 : 데이터를 바탕으로 기술적 지표(이동 평균선, 변동성)와 재무적 지표(영업이익률, 순이익률)를 계산하여 모델 학습의 힌트(Feature)를 만듦

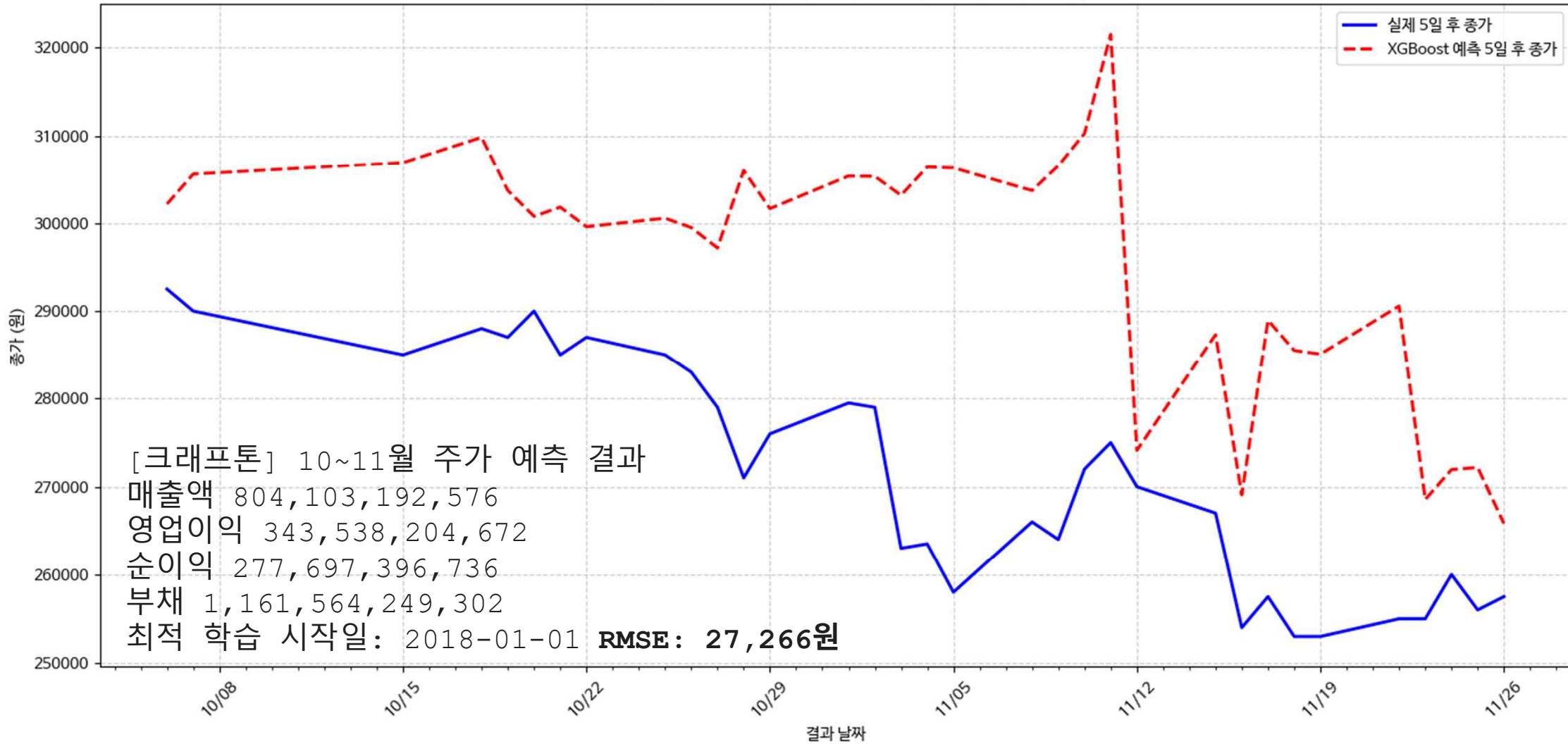
### 2. 1단계 예측 : 미래 재무제표 예측(LSTM)

- 예측을 할 시점의 실적이 아직 발표되지 않았으므로, LSTM을 이용해 4분기의 재무제표를 우선 예측.

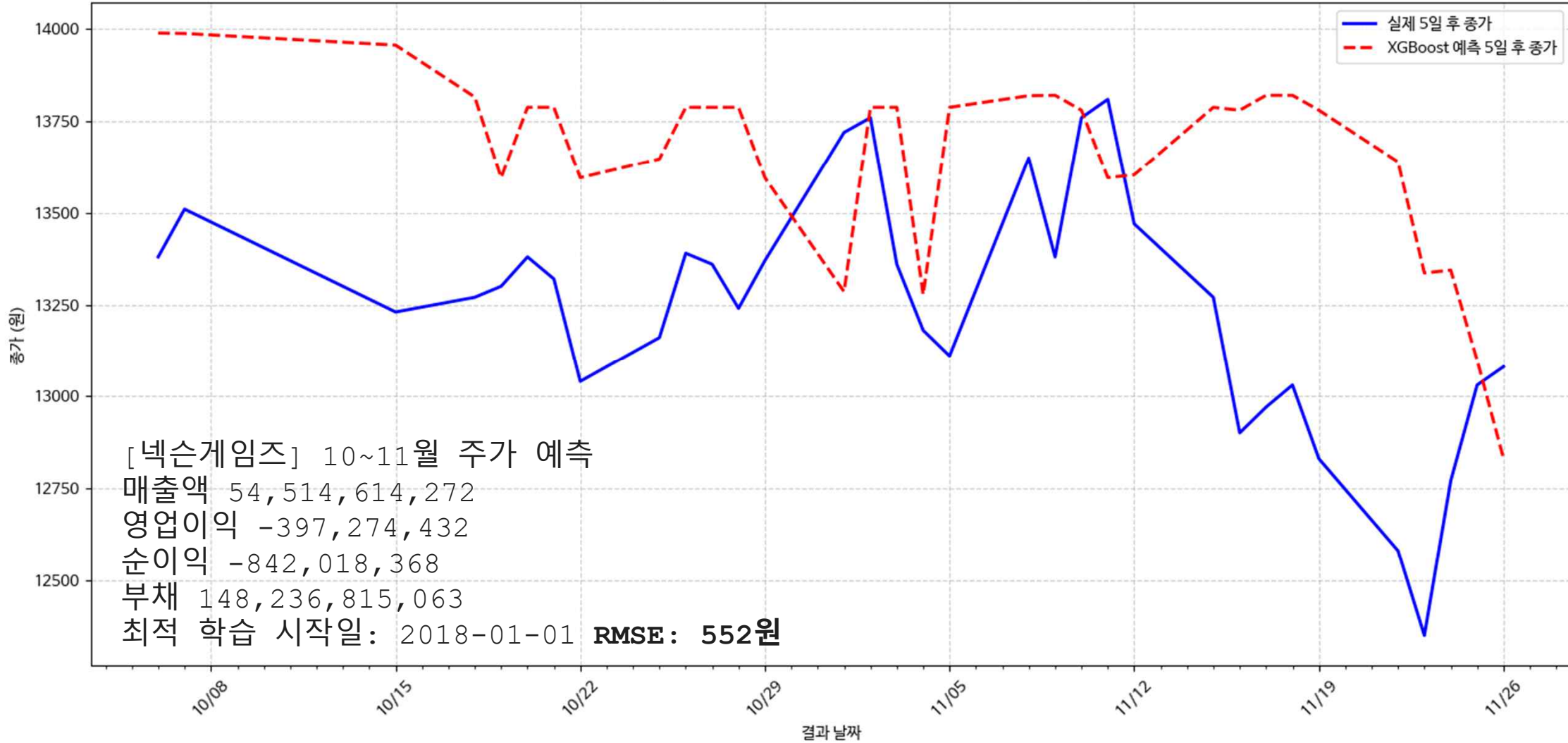
### 3. 2단계 예측 : 주가 예측(XGBoost)

- 2단계에서 예측한 4분기 실적과 과거 주가 패턴을 합쳐서 XGBoost 모델이 최종 주가를 예측
- 과거 주가 흐름과 재무 상태가 주가에 미친 영향을 학습.
- GridSearchCV를 통해 가장 성능이 좋은 최적의 파라미터와 학습 시작 기간을 찾아냄.

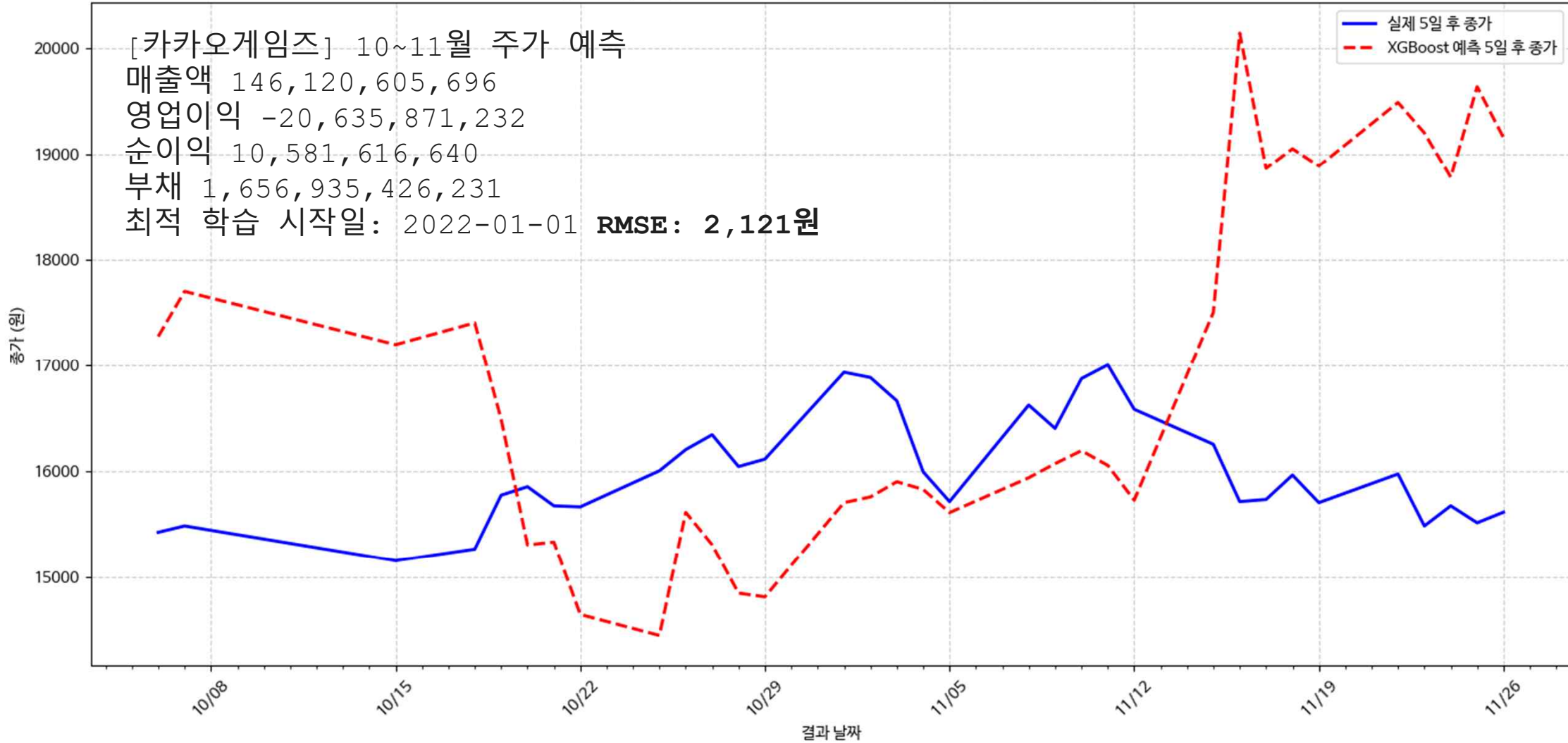
[크래프톤] XGBoost 5일 후 주가 예측 비교 (2025년 10월~11월)



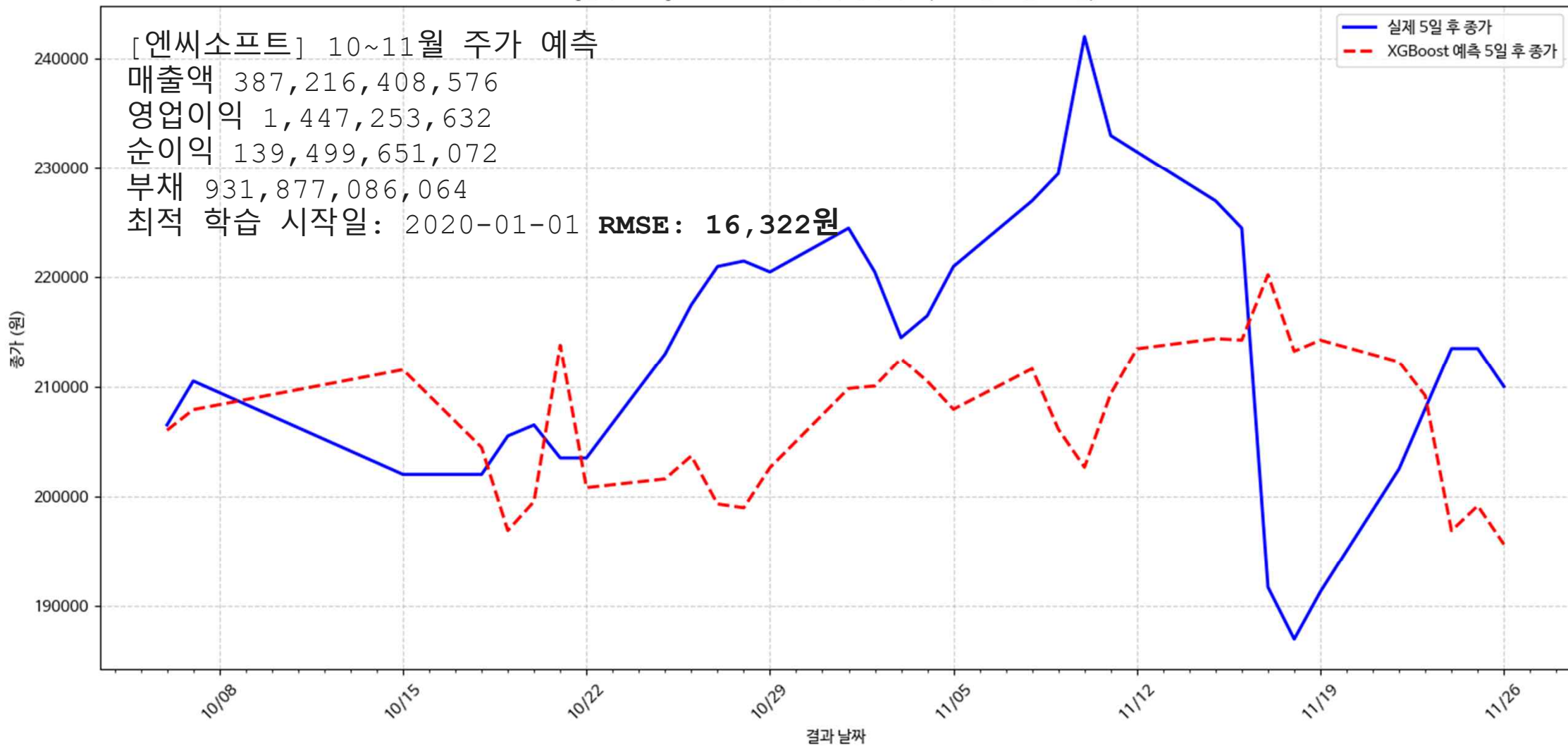
[넥슨게임즈] XGBoost 5일 후 주가 예측 비교 (2025년 10월~11월)



[카카오게임즈] XGBoost 5일 후 주가 예측 비교 (2025년 10월~11월)

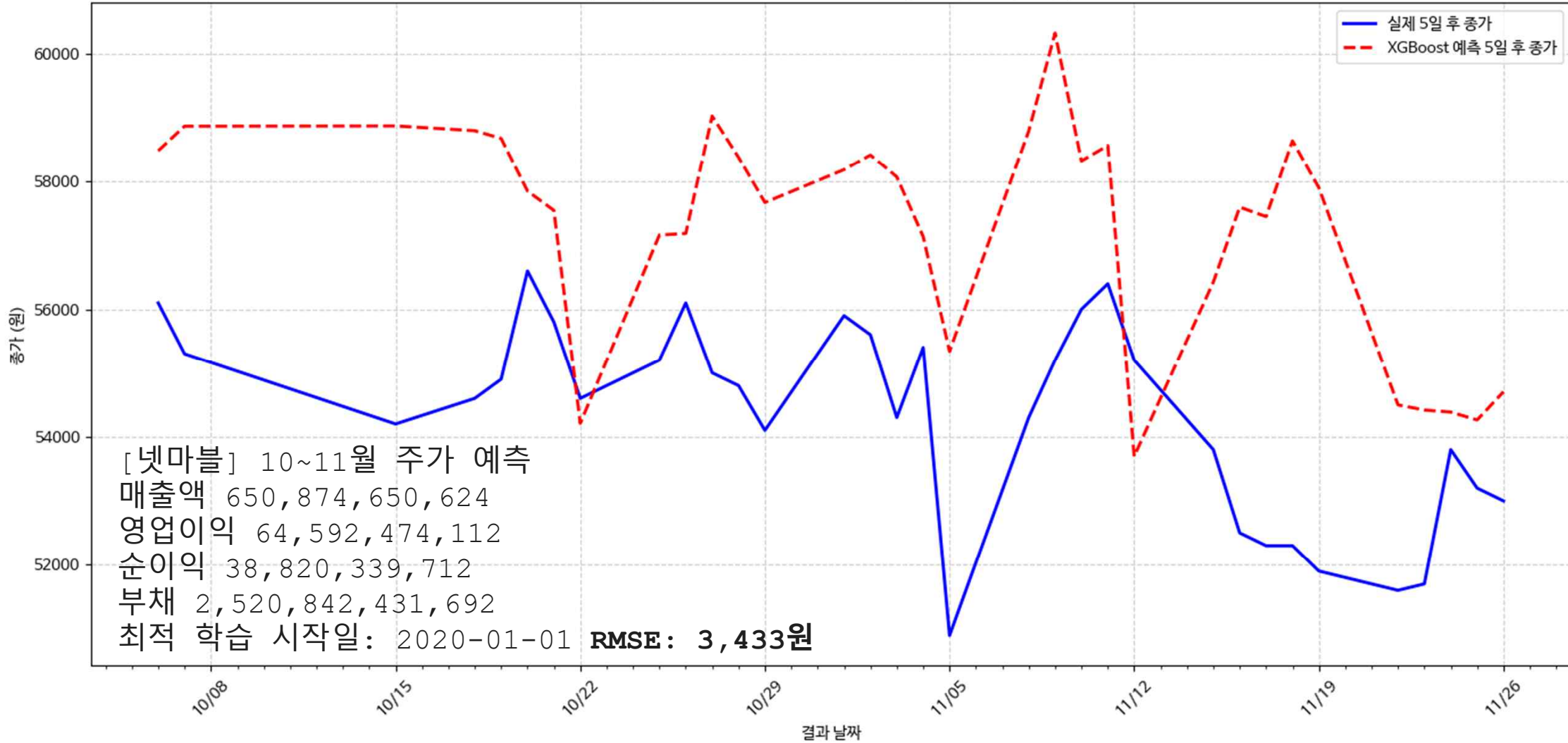


[엔씨소프트] XGBoost 5일 후 주가 예측 비교 (2025년 10월~11월)





[넷마블] XGBoost 5일 후 주가 예측 비교 (2025년 10월~11월)



---

# CONCLUSION

**모델의 구조적 한계**: XGBoost와 같은 트리 기반 회귀 모델은 예측 값과 실제 값 사이에 시차(Lag)가 발생하거나, 급격한 변동을 따라가지 못하는 과리가 관측됨.

- WHY?**
- **Memory 부재**: RNN이나 LSTM과 달리, XGBoost는 이전 시점의 상태(State)를 기억하는 **순환 구조**가 없다. 또한, 과거 데이터를 독립적인 변수로만 처리하기 때문에 시계열적 연속성을 완벽히 포착하기 어렵다.
  - **보수적 예측**: 회귀 모델은 학습 데이터의 **평균적인 패턴**을 학습하려는 경향이 강함. 따라서 주가 폭등이나 폭락 같은 **이상치 성격의 움직임**보다는, 안정적인 평균값으로 수렴하려는 예측 결과를 내놓아 실제 주가 변동성보다 **밋밋한 예측선**이 그려지게 됨.
-

---

# CONCLUSION

**정보의 비대칭성:** 재무 상태(매출, 영업이익)를 학습시켰음에도 불구하고, 실제 주가는 재무 데이터와 무관하게 움직이거나 예측력을 벗어나는 경우가 많았음.

**WHY? • 선반영 효과:** 주식 시장은 효율적 시장 가설에 따라 공개된 재무 정보는 이미 주가에 반영되어있다. 주가는 과거가 아닌 **미래의 기대**에 의해 움직이기 때문에, 과거 재무 데이터의 예측 기여도는 제한적일 수밖에 없었음.

- **외부 변수의 지배력:** 현재 게임 업계는 기업 내부의 실적보다 금리, 물가, 정부 규제 등 거시경제 변수와 신작 출시 지연 같은 외부 뉴스(Event)에 훨씬 민감하게 반응하고있음. 이는 정형 데이터만으로는 예측할 수 없는 영역임
-

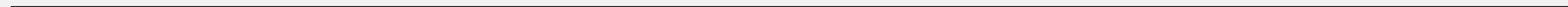
---

# CONCLUSION

**패턴의 규칙성**: 크래프톤이나 엔씨소프트처럼 추세가 급변하는 종목보다, 넥슨게임즈와 같이 일정 범위 내에서 등락을 반복하는 종목의 RMSE가 현저히 낮게 측정되었습니다

**WHY?** • **데이터의 정상성**: 머신러닝 모델은 학습 데이터의 분포가 일정할 때(패턴이 반복될 때) 가장 높은 성능을 발휘하게 됨.

- **오실레이터 패턴**: 주가가 올랐다가 다시 평균으로 돌아오는 패턴이 반복되는 종목은 모델이 "이쯤 오르면 내리고, 이쯤 내리면 오른다"는 규칙을 학습하기 매우 유리함. 반면, 끝없이 하락하거나 상승하는 추세장에서는 모델이 경험하지 못한 가격대라 예측력이 급격히 떨어지게 됨.



# 라이브러리

---

## 데이터 수집 및 처리

`pykrx` : 한국 거래소(KRX) 주가 데이터(OHLCV) 확보용

`requests` : API 쿼리 요청 (재무 데이터 수집 등 외부 통신용)

`pandas` : 시계열 주가 및 재무 데이터 프레임 생성, 병합 (`merge_asof`), 전처리

`numpy` : 수치 연산, 배열 처리, 결측치 대응 및 `log/exp` 변환용

`zipfile, io` : 압축 파일 입출력 및 메모리 내 파일 처리

`xml.etree.ElementTree` : XML 형식의 재무 데이터 파싱

---

---

# 머신러닝 및 딥러닝 모델링

`xgboost` (XGBRegressor) : 주가 예측을 위한 메인 회귀 모델  
(Gradient Boosting 알고리즘)

`tensorflow.keras` (LSTM, Dense, Sequential) : 재무 항목 시  
계열 예측을 위한 딥러닝 모델

`sklearn.preprocessing` (MinMaxScaler) : 데이터 정규화 (0~1  
사이로 스케일링)

`sklearn.model_selection` (GridSearchCV, TimeSeriesSplit) :  
최적 하이퍼파라미터 탐색 및 시계열 교차 검증

`sklearn.metrics` (mean\_squared\_error) : 모델 성능 평가  
(RMSE 계산용)

---

---

# 시각화 및 유틸리티

`matplotlib.pyplot`, `matplotlib.dates` : 예측 결과 및 주가 그래프 시각화, 날짜 포매팅

`plotly` : 인터랙티브 데이터 시각화 (동적 그래프 필요 시 사용)

`math` : RMSE 계산(`sqrt`) 등 수학 함수 지원

`os`, `sys` : 파일 경로 제어, 시스템 설정 및 예외 처리

`warnings` : 불필요한 경고 메시지 제어 (출력 깔끔하게 유지)

---



THANK YOU

뉴하트

