

Results

Statistical Insights

Panel A. Summary Statistics							
	Mean	SD	Min	P25	Median	P75	Max
Daily Return(%)	0.017	0.517	-7.011	-0.25	0.000	0.244	2.396
Headline Length	32.047	8.454	2.0	26.0	32.0	37.0	109.0
GPT4.1Nano Score	0.213	0.788	-1.000	-0.0	0.000	1.000	1.000

- **Portfolio Returns** reflect only days when positions were actually taken, capturing the effective returns of the strategy on active trading days.
- **Panel A** calculates average returns for **all news headlines**, meaning erroneous positions can significantly lower overall average returns.
- **Daily Return (%)**:
 - Average daily return: approximately **0.017%**, close to zero.
 - High volatility (standard deviation exceeding $\pm 0.5\%$), with returns ranging widely from **7% to +2.39%**.
- **Headline Length**:
 - Median headline length: 32 characters, with most headlines between 26 and 37 characters.
 - Maximum headline length is 109 characters, indicating some extremely long headlines.
- **GPT Score**:
 - Mean score: 0.213; median: 0.0, indicating a slight overall positive bias.
- Distribution of returns shows significant skewness and extreme values, suggesting sensitivity of investment returns to specific market events.

Panel B. Correlations			
	Daily Return(%)	Headline Length	GPT4.1Nano Score
Daily Return(%)	1.000	-0.002	0.015
Headline Length	-0.002	1.000	-0.021
GPT4.1Nano Score	0.015	-0.021	1.000

Correlation Analysis Summary

We analyzed relationships among the GPT4.1Nano sentiment scores, headline lengths, and daily returns. The resulting Pearson correlation coefficients were very weak (absolute values below 0.02):

- GPT score \leftrightarrow Daily Return: **+0.015**
- Headline Length \leftrightarrow Daily Return: **0.002**
- GPT score \leftrightarrow Headline Length: **0.021**

These results imply that the GPT sentiment scores do **not exhibit a direct linear relationship** with either returns or structural characteristics of news headlines.

Reasons for Weak Correlation

Previous studies employing GPT-3 or GPT-3.5 found more pronounced correlations between sentiment scores and returns due to limitations in earlier language models, which primarily captured surface-level sentiment.

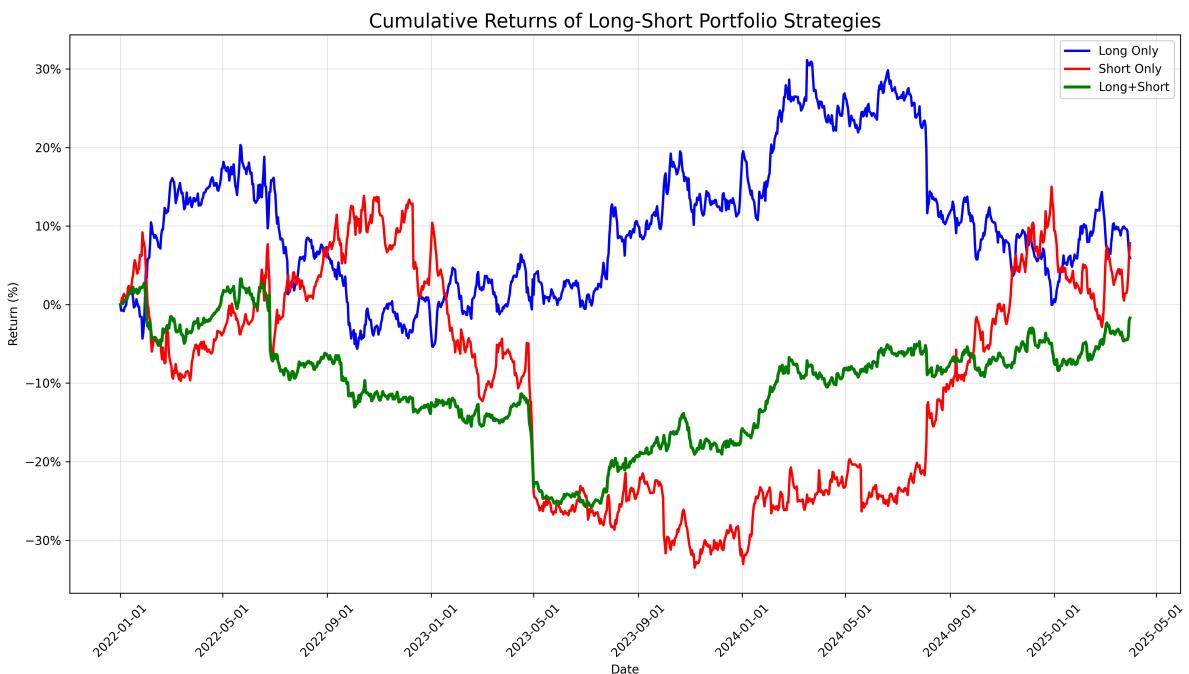
However, this study utilized the **GPT-4.1 Nano** model, characterized by:

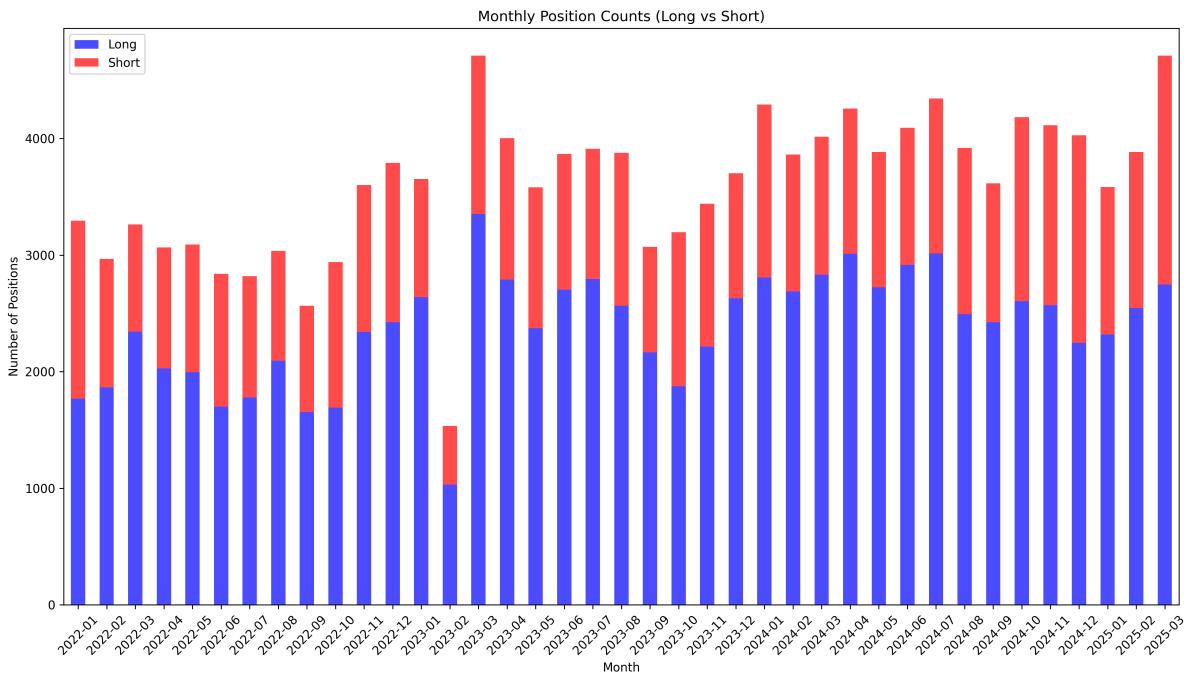
- Significantly enhanced contextual understanding and interpretation capabilities.
- More nuanced and neutral sentiment assessments rather than simplistic positive/negative categorizations.

Conclusion and Implications

- While GPT-4.1 Nano substantially improved **textual analysis quality**, its sophisticated interpretative capabilities resulted in a **weaker simple correlation** between sentiment scores and investment returns.
- These findings highlight that enhanced language models, though superior in understanding textual nuances, may lead to reduced effectiveness when using direct sentiment-based approaches for predicting financial returns.

Total Cumulative Return





Summary Statistics:

- Total news analyzed:** 270,426
- Long positions:** 88,850
- Short positions:** 46,006

Annualized Returns (based on actual trading days)

Strategy	Annualized Return (%)
Long-only	1.79%
Short-only	2.35%
Long-Short Combined	-0.54%

Reason for Negative Long-Short Combined Return

The combined long-short strategy calculates returns using a weighted average of daily long and short positions rather than a simple average. As a result, positions became heavily weighted toward the losing side (either long or short) during certain periods, causing overall negative returns.



News Based Result

Data Shape: (270426, 23)

News Distribution by Category

No.	Category (통합 분류1)	Number of Articles
1	Economy > Stock Market	67,471
2	Economy > Industry & Corporates	36,693
3	Economy > Retail	31,799
4	Economy > Automotive	24,547
5	Economy > Real Estate	22,334
6	Economy > Finance & Investment	20,537
7	Economy > Semiconductors	14,677
8	Economy > Services & Shopping	13,530
9	Economy > General Economy	12,595
10	Economy > Employment & Startups	7,201
11	Economy > Resources	6,564
12	Economy > Trade	5,325
13	Economy > Global Economy	4,616
14	Economy > Foreign Exchange	253

Average Daily Number of Traded Stocks by News Category

No.	Category (통합 분류1)	Avg. Daily Traded Stocks
1	Economy > Stock Market	56.89
2	Economy > Industry & Corporates	30.94
3	Economy > Retail	26.81
4	Economy > Automotive	20.70
5	Economy > Real Estate	18.83
6	Economy > Finance & Investment	17.32
7	Economy > Semiconductors	12.38
8	Economy > Services & Shopping	11.41
9	Economy > General Economy	10.62
10	Economy > Employment & Startups	6.07
11	Economy > Resources	5.53

12	Economy > Trade	4.49
13	Economy > Global Economy	3.89
14	Economy > Foreign Exchange	2.14

Performance Metrics by News Category

No.	Category (통합 분류1)	Annualized Return	Annualized Volatility	Sharpe Ratio	Max Drawdown
1	Economy > General Economy	15.59%	17.22%	0.91	-12.66%
2	Economy > Global Economy	-2.29%	10.42%	-0.22	-33.64%
3	Economy > Finance & Investment	2.73%	7.47%	0.37	-10.10%
4	Economy > Trade	3.15%	11.56%	0.27	-22.48%
5	Economy > Semiconductors	-3.08%	11.54%	-0.27	-36.64%
6	Economy > Real Estate	-4.44%	6.77%	-0.66	-23.04%
7	Economy > Industry & Corporates	1.20%	9.62%	0.12	-18.17%
8	Economy > Services & Shopping	-4.57%	8.08%	-0.57	-21.51%
9	Economy > Foreign Exchange	0.87%	9.02%	0.10	-14.89%
10	Economy > Retail	1.45%	8.95%	0.16	-14.23%
11	Economy > Automotive	-5.67%	9.78%	-0.58	-30.36%
12	Economy > Resources	-16.14%	18.79%	-0.86	-68.71%
13	Economy > Stock Market	3.45%	10.53%	0.33	-22.76%
14	Economy > Employment & Startups	3.49%	8.12%	0.43	-16.86%
15	Overall	-0.53%	3.23%	-0.17	-6.25%

Performance Analysis by Time Period

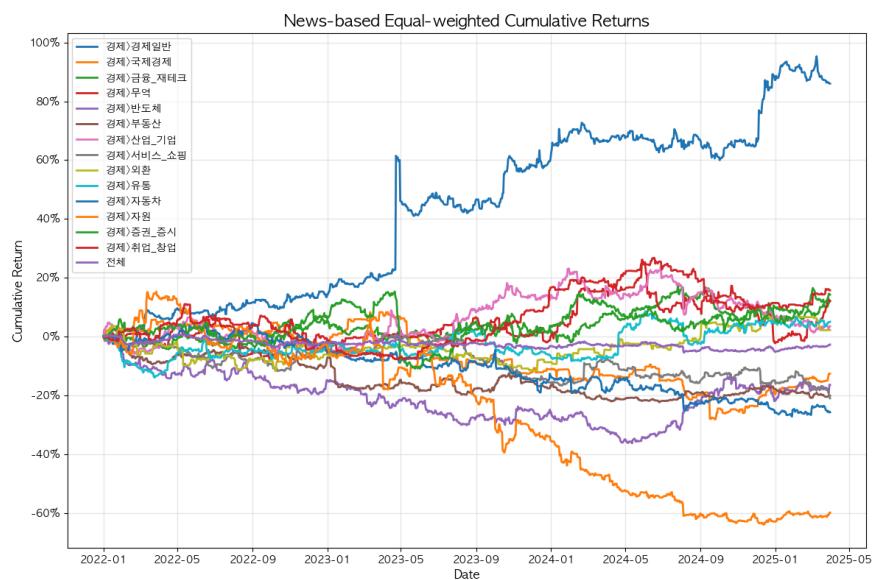
Period	Annualized Return	Volatility (Annualized)	Sharpe Ratio	Max Drawdown
2022 (Full Year)	1.92%	4.10%	0.47	-5.84%
2023 (Full Year)	-1.76%	3.95%	-0.45	-7.12%
2024 (Jan–Dec)	0.38%	3.21%	0.12	-4.67%
2025 (Jan–Mar)	-3.12%	2.95%	-1.06	-3.79%

Cumulative, Monthly, and Quarterly Returns by News Category

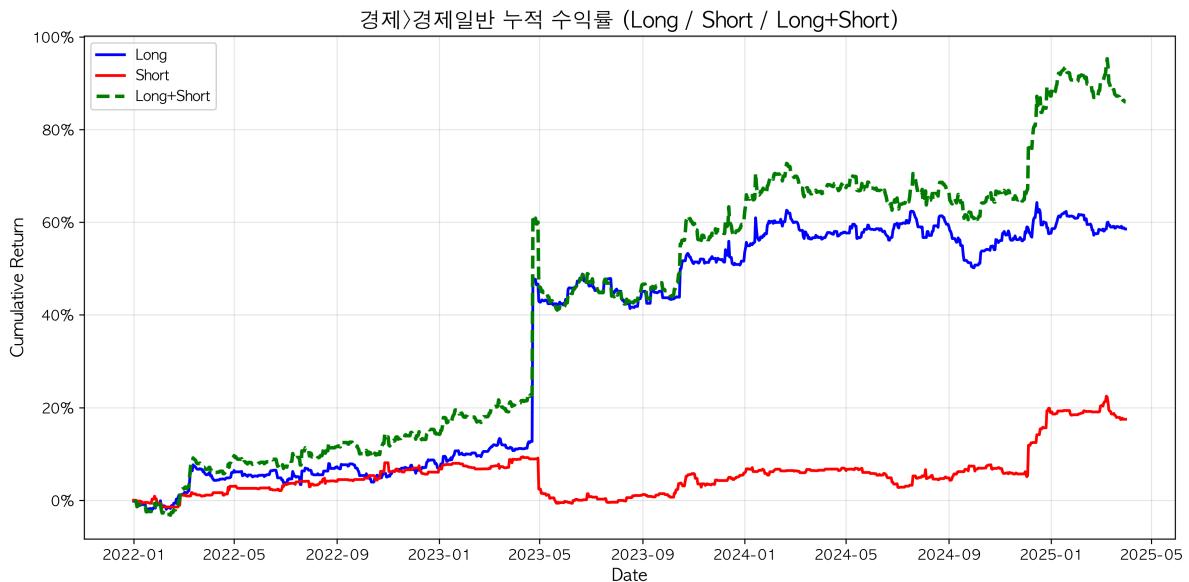
No.	Category (통합 분류1)	Total Return (%)	Avg. Monthly Return (%)	Avg. Quarterly Return (%)
1	Economy > General Economy	86.02%	1.70%	5.12%
2	Economy > Global Economy	-12.63%	-0.29%	-0.77%
3	Economy > Finance & Investment	12.02%	0.33%	0.92%
4	Economy > Trade	12.15%	0.42%	1.28%
5	Economy > Semiconductors	-16.37%	-0.40%	-1.07%
6	Economy > Real Estate	-20.12%	-0.54%	-1.63%
7	Economy > Industry & Corporates	3.49%	0.14%	0.42%
8	Economy > Services & Shopping	-21.00%	-0.56%	-1.65%
9	Economy > Foreign Exchange	2.18%	0.09%	0.22%
10	Economy > Retail	5.03%	0.18%	0.55%
11	Economy > Automotive	-25.74%	-0.72%	-2.21%
12	Economy > Resources	-59.87%	-2.16%	-6.31%

13	Economy > Stock Market	14.27%	0.44%	1.28%
14	Economy > Employment & Startups	15.70%	0.41%	1.28%
15	Overall	-2.73%	-0.07%	-0.20%

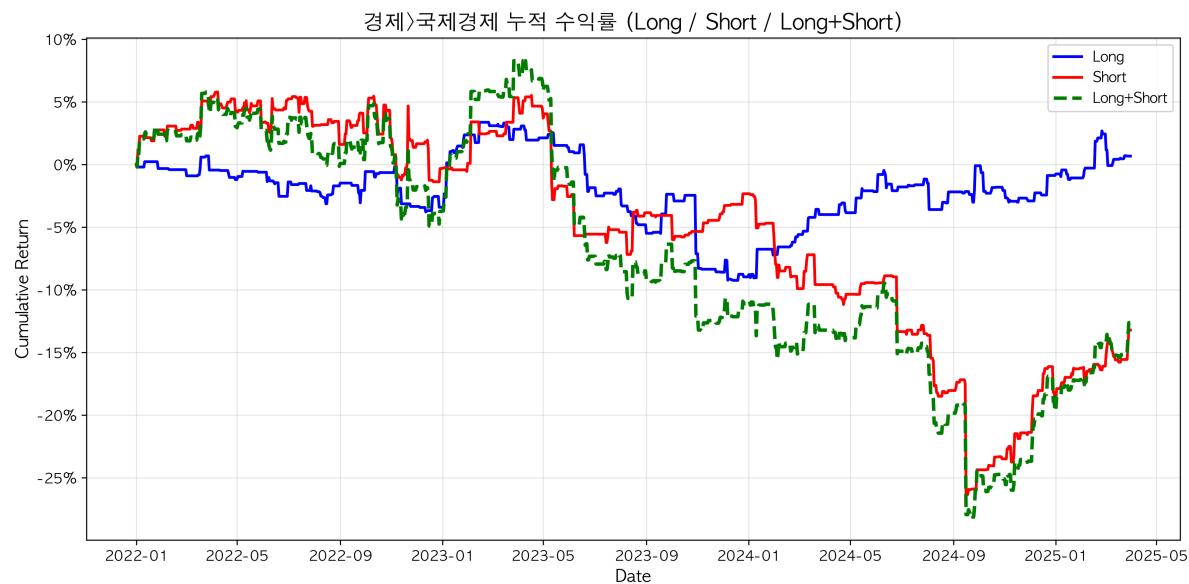
New-based



General Economy



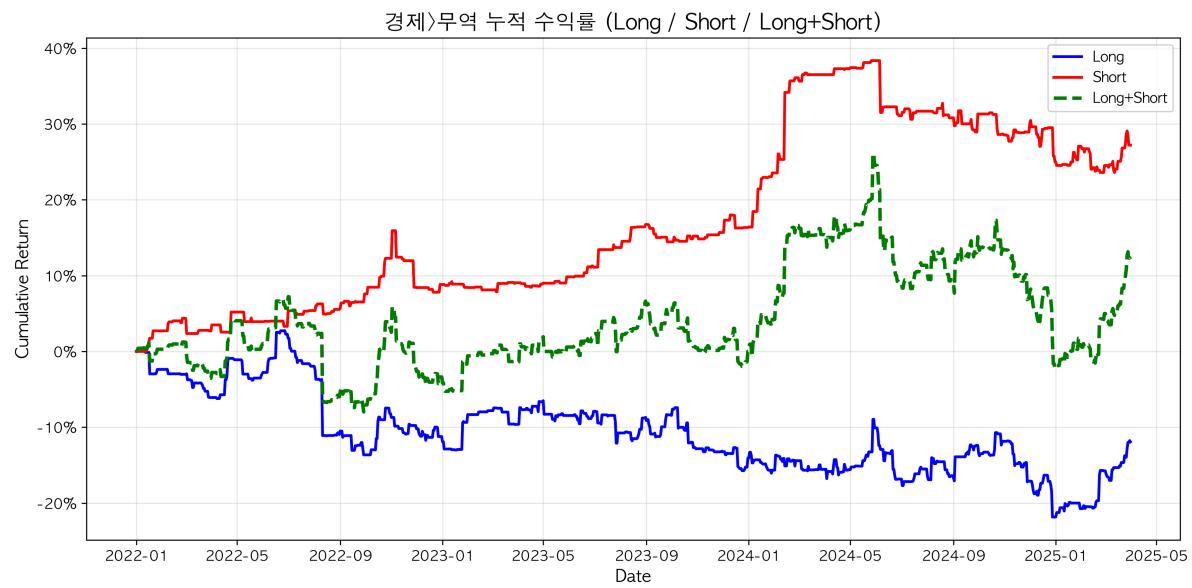
Global Economy



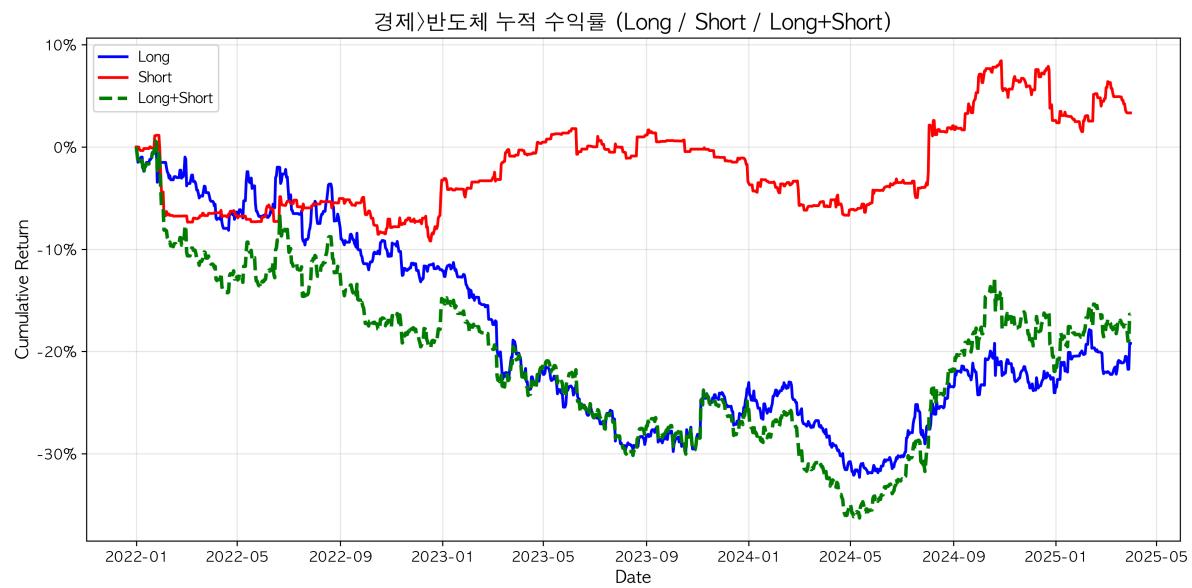
Finance & Investment



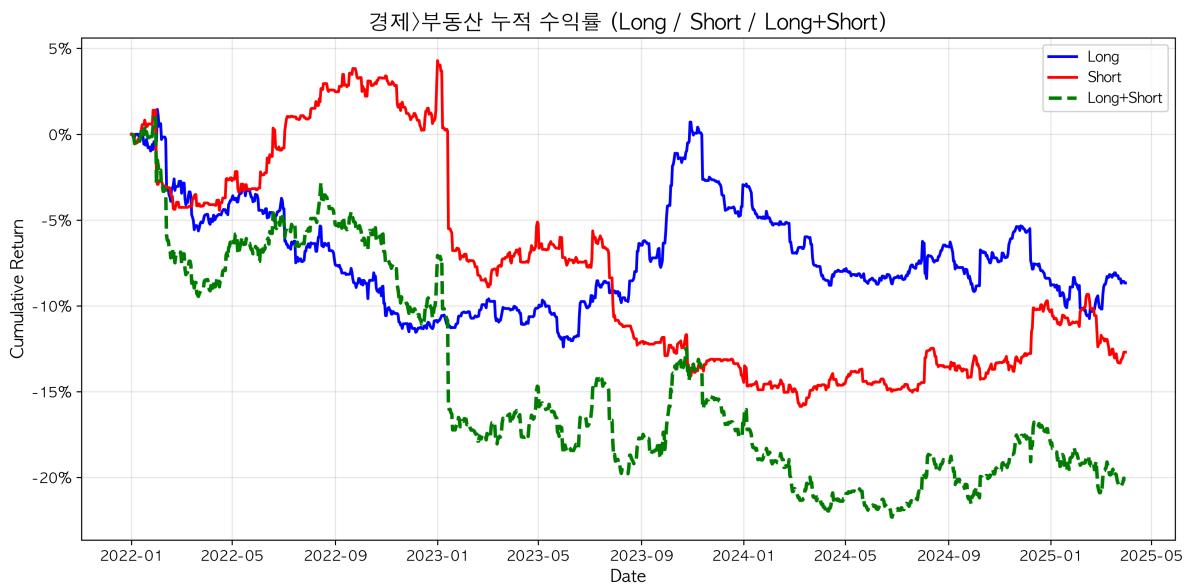
Trade



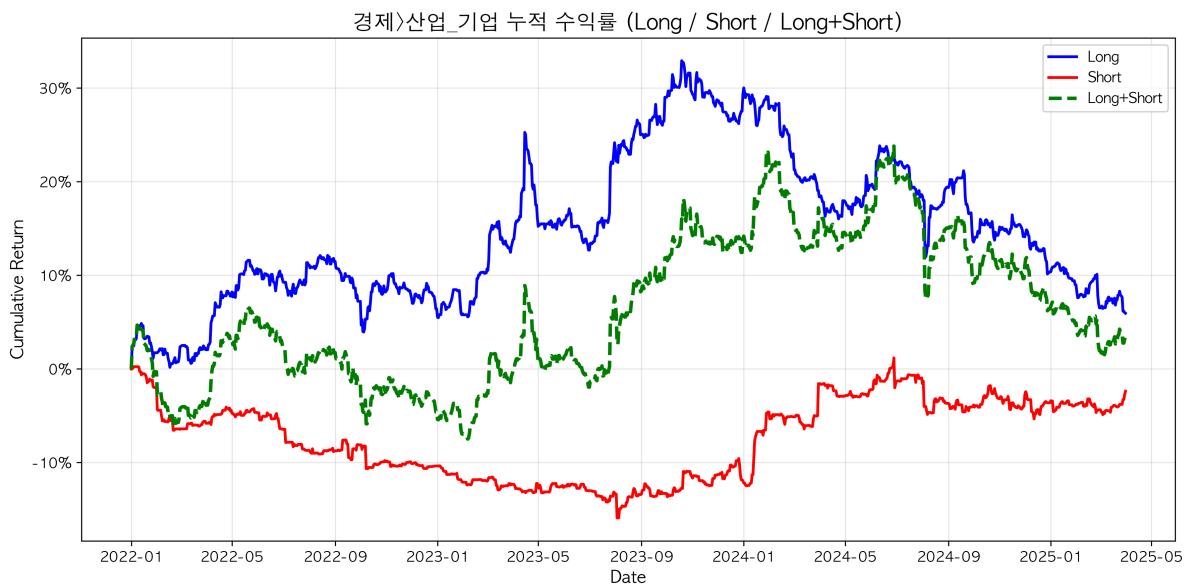
Semiconductors



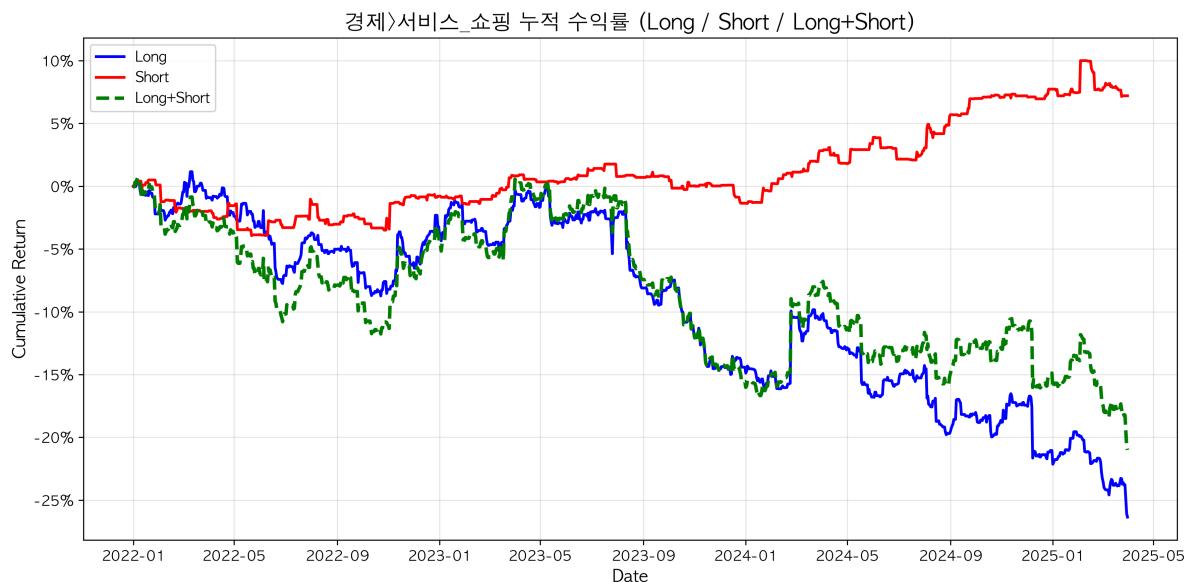
Real Estate



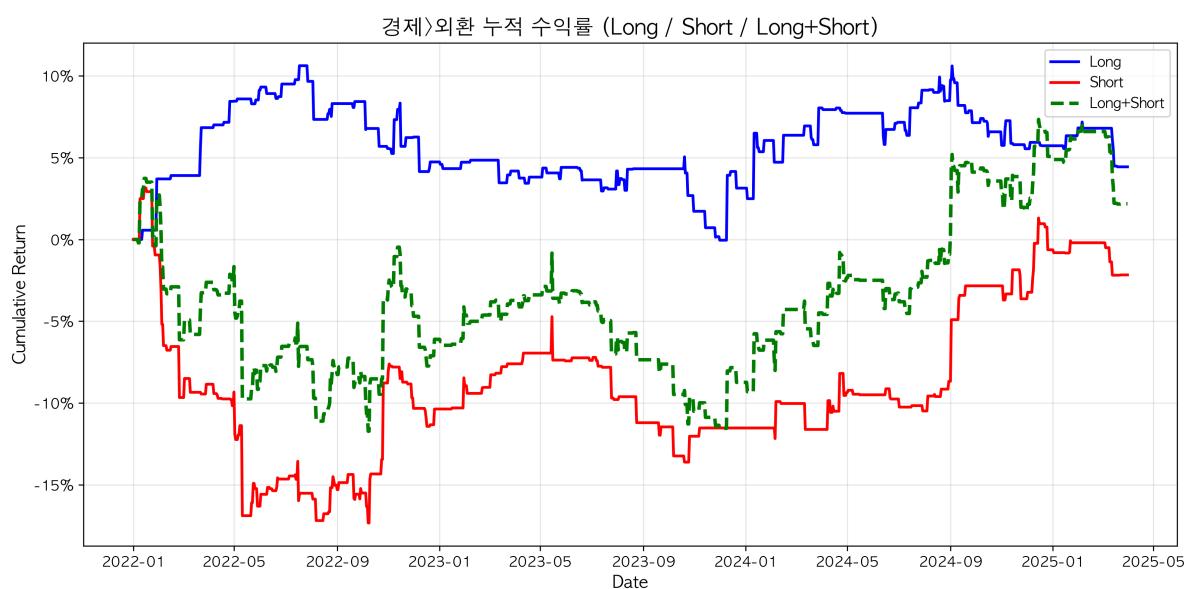
Industry & Corporates



Services & Shopping

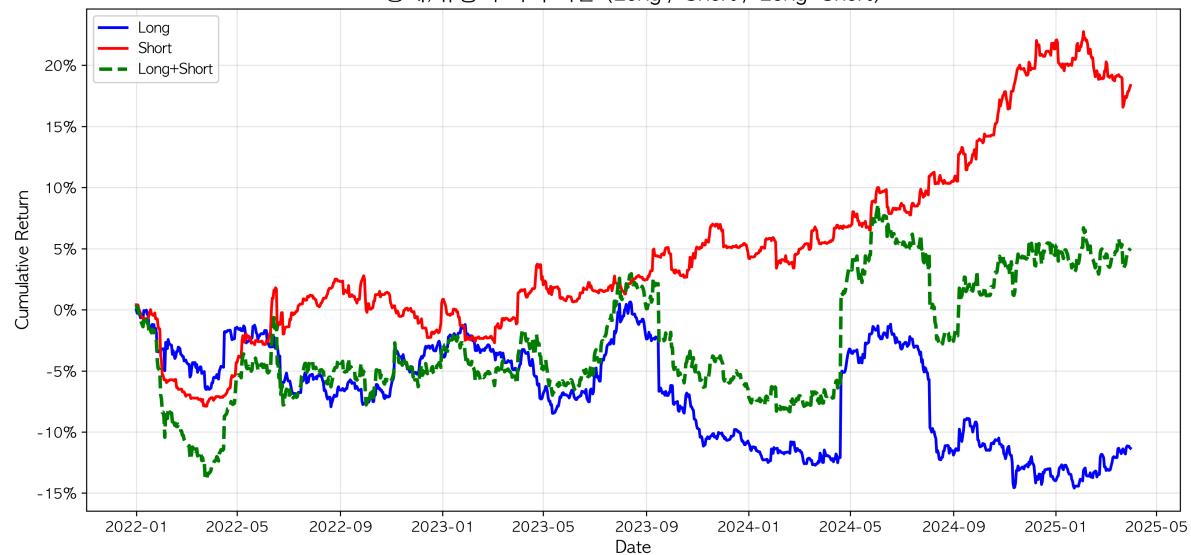


Foreign Exchange



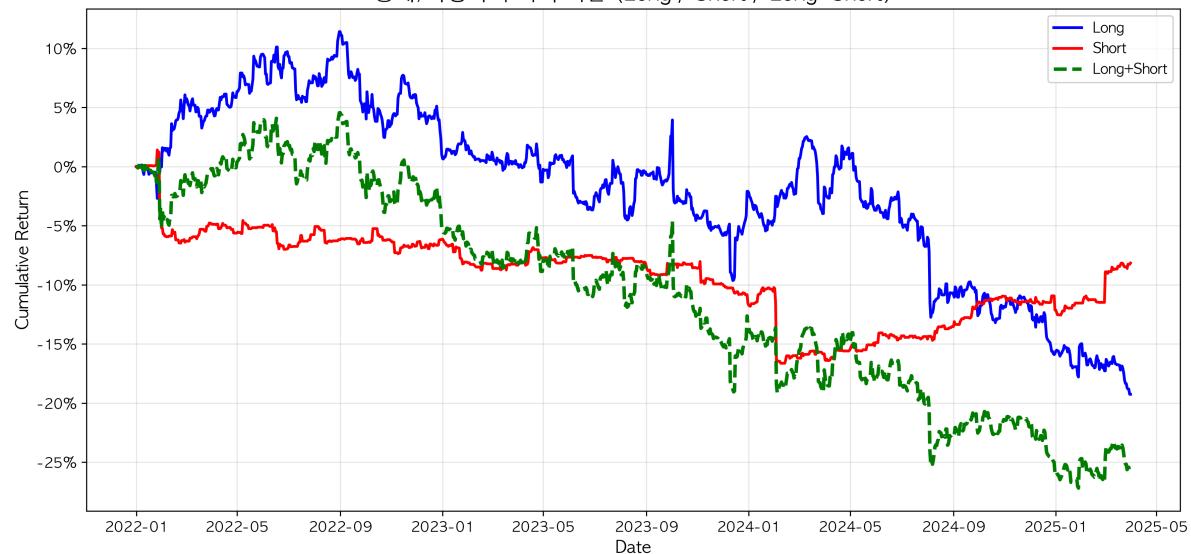
Retail

경제>유통 누적 수익률 (Long / Short / Long+Short)

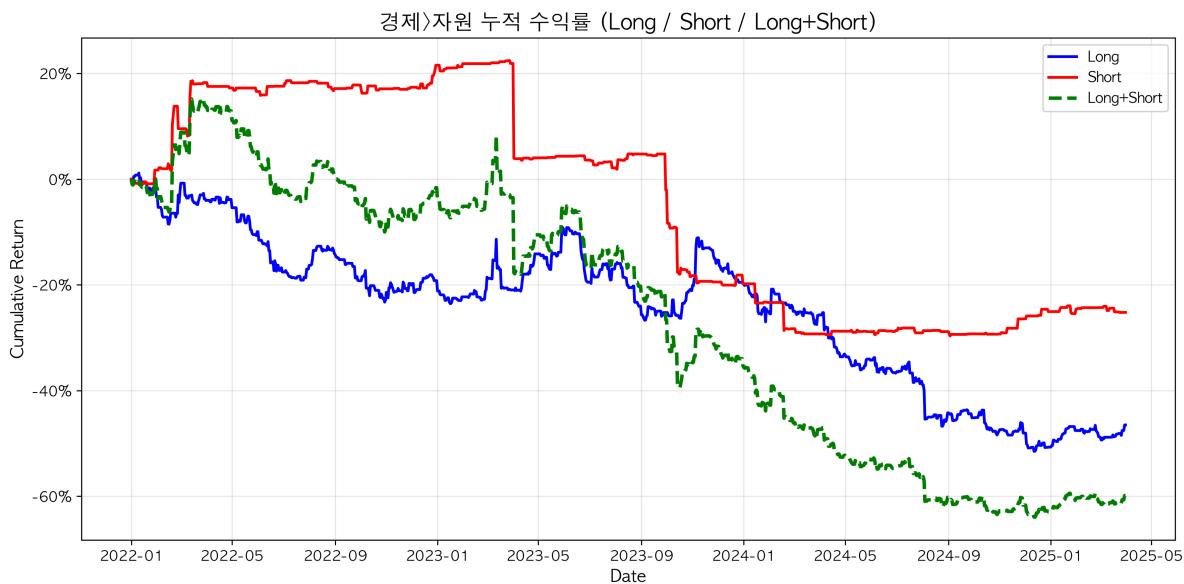


Automotive

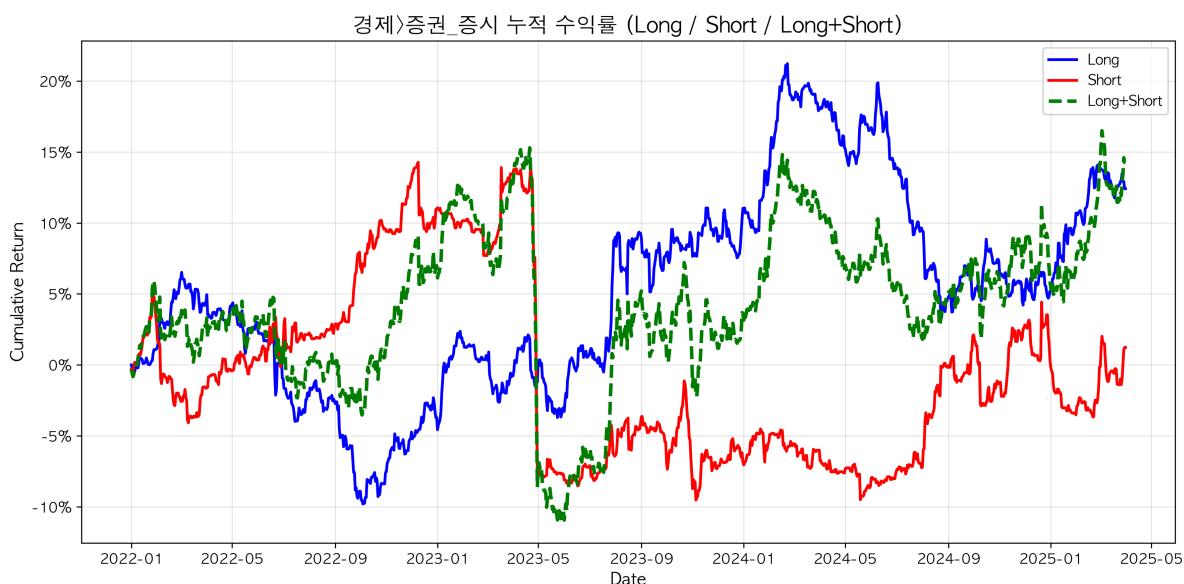
경제>자동차 누적 수익률 (Long / Short / Long+Short)



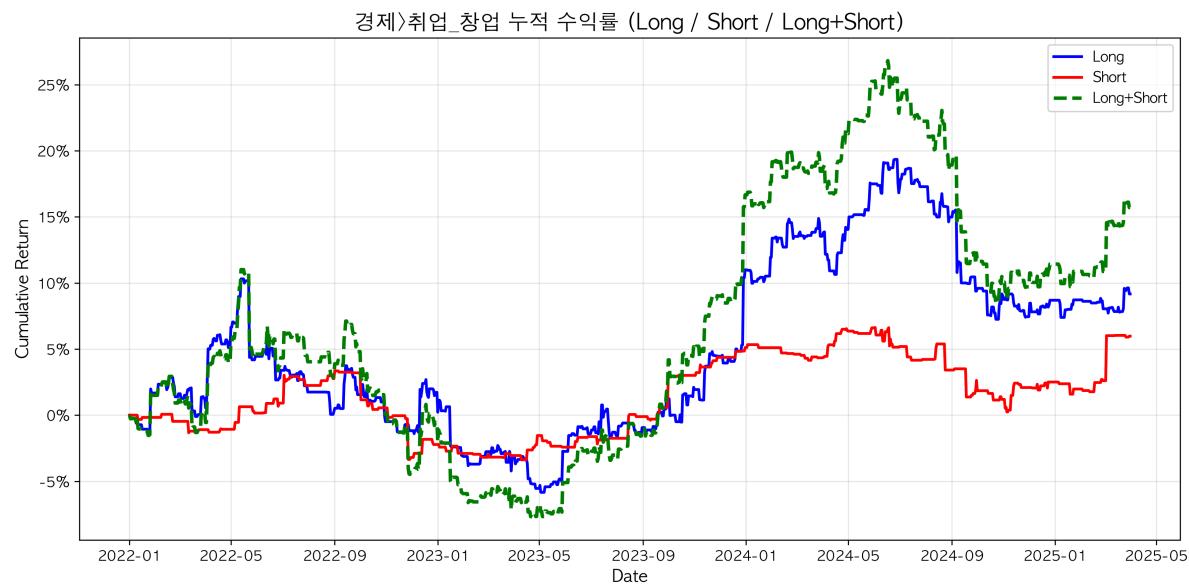
Resources



Stock Market



Employment & Startups



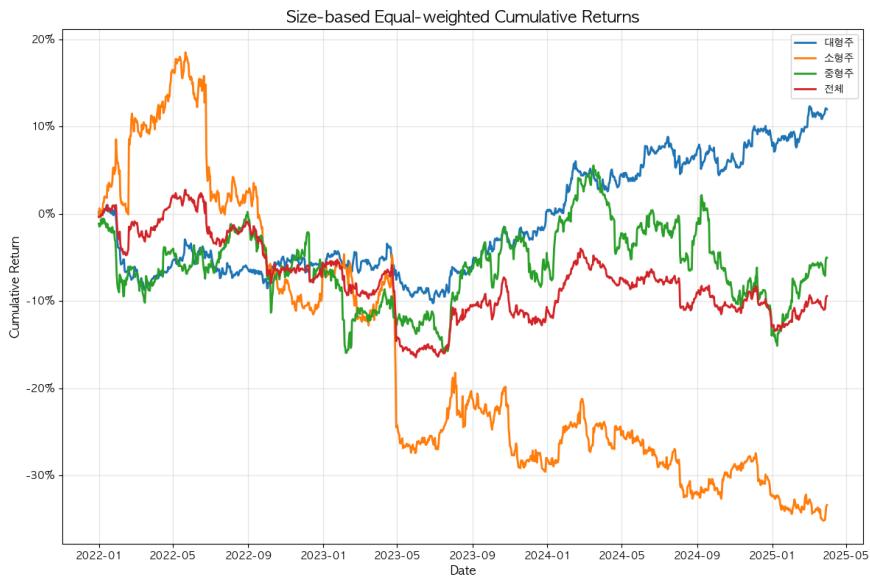
Size-based

No.	Firm Size Category	Number of News Articles	Avg. Daily Traded Stocks
1	Large-cap	126,421	106.68
2	Small-cap	48,769	41.16
3	Mid-cap	31,218	26.34

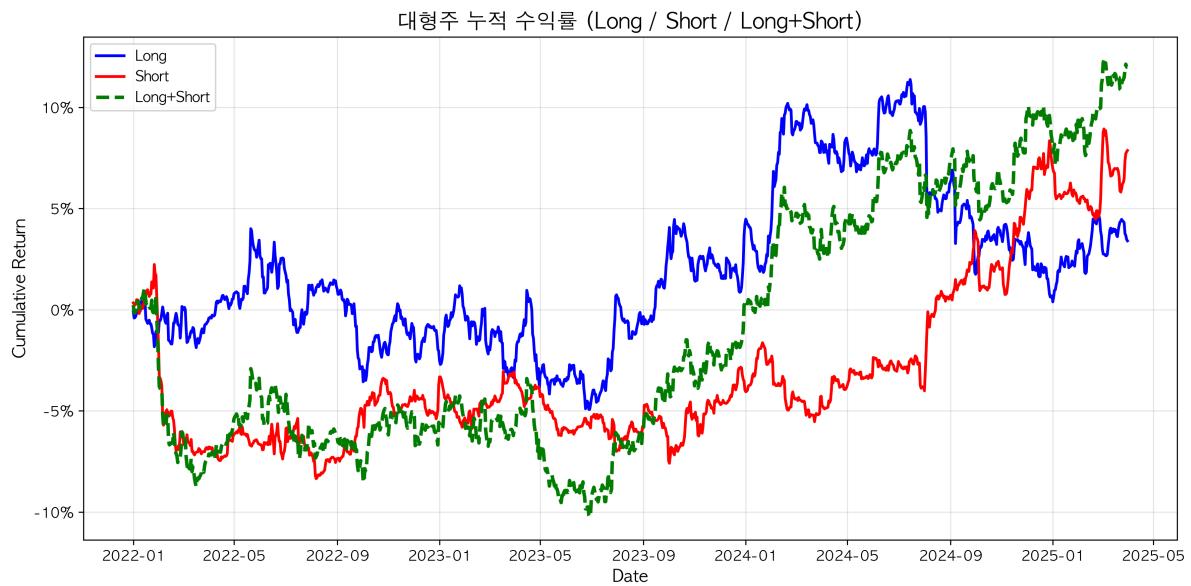
No.	Firm Size Category	Annualized Return	Annualized Volatility	Sharpe Ratio	Max Drawdown
1	Large-cap	2.62%	5.89%	0.44	-11.16%
2	Small-cap	-7.45%	13.37%	-0.56	-45.31%
3	Mid-cap	-0.50%	10.86%	-0.05	-19.59%
4	Overall	-1.87%	6.56%	-0.29	-18.71%

No.	Firm Size Category	Total Return (%)	Avg. Monthly Return (%)	Avg. Quarterly Return (%)
1	Large-cap	11.99%	0.31%	0.95%
2	Small-cap	-33.39%	-0.93%	-2.84%
3	Mid-cap	-5.01%	-0.09%	-0.23%
4	Overall	-9.42%	-0.23%	-0.68%

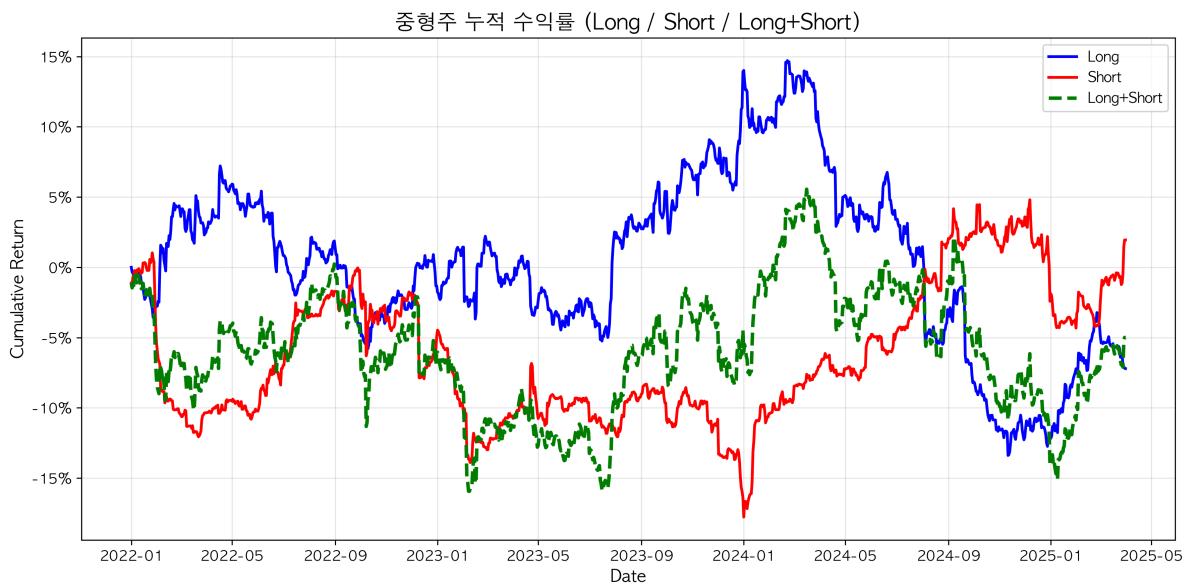
Total Result



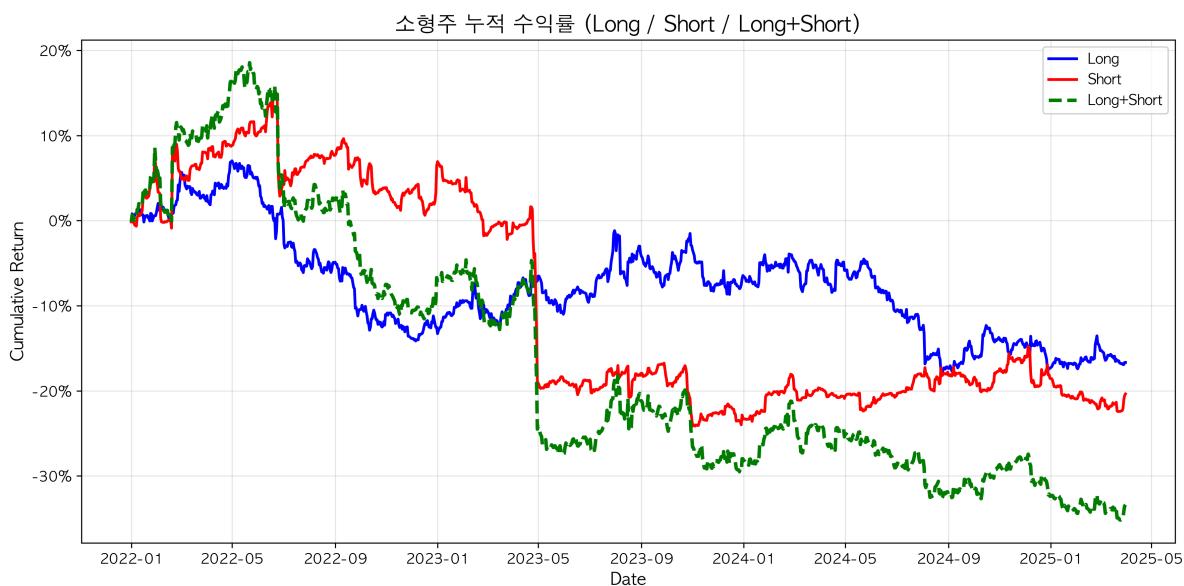
Large



Mid



Small



Style-based

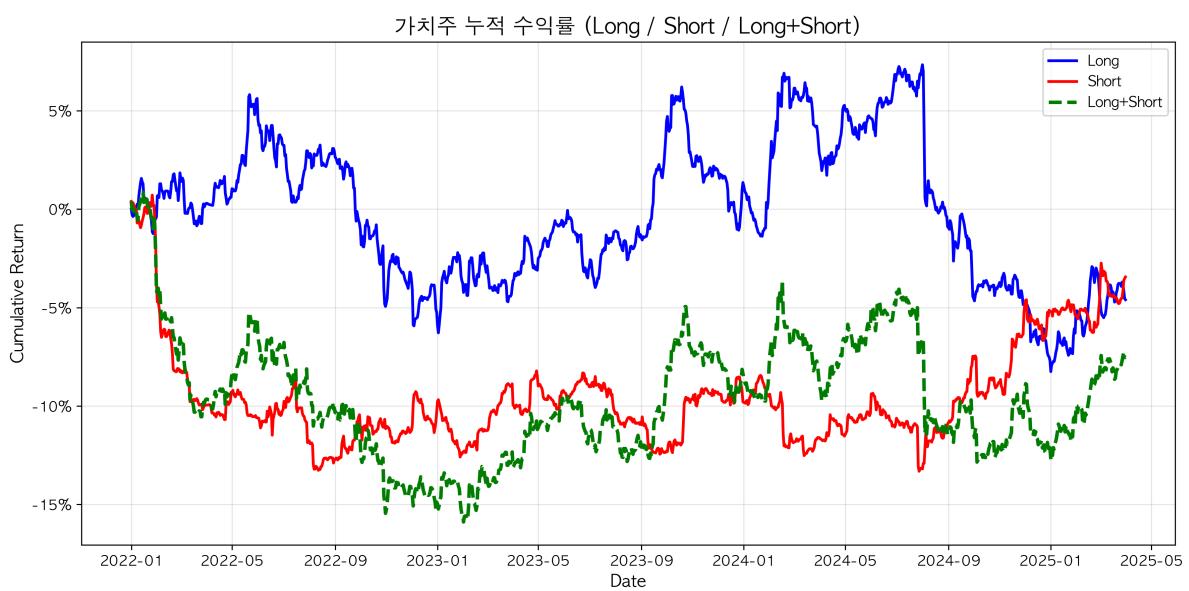
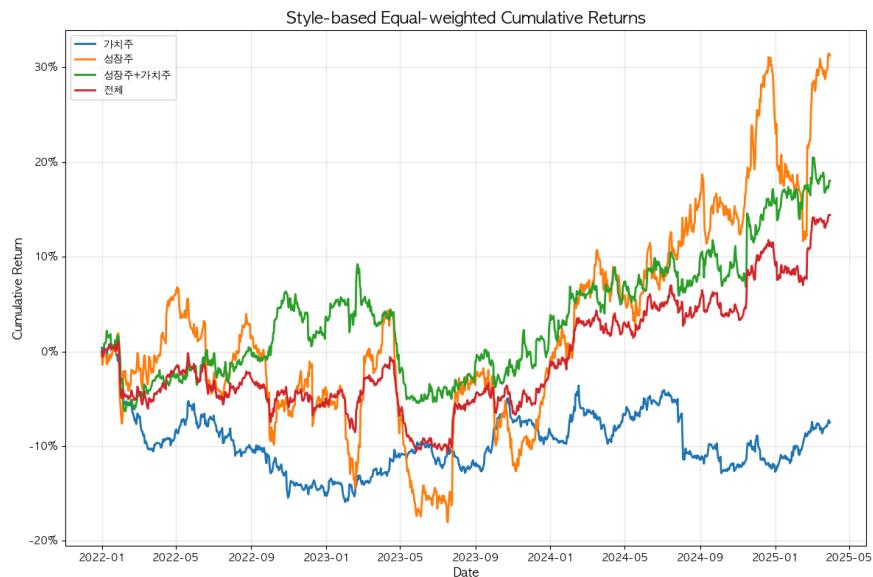
Kodex classification

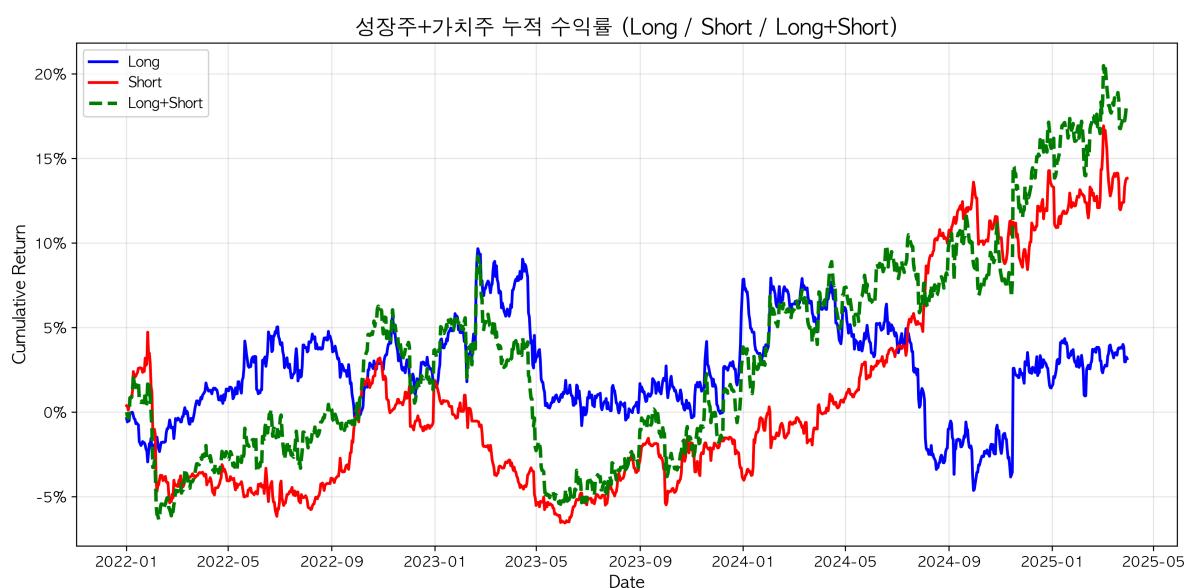
No.	Investment Style	Number of Stocks Mentioned in News	Avg. Daily Traded Stocks
1	Value Stocks	59,320	50.19
2	Growth + Value Stocks	47,585	40.26
3	Growth Stocks	29,041	24.57

No.	Investment Style	Annualized Return	Annualized Volatility	Sharpe Ratio	Max Drawdown
1	Value Stocks	-1.47%	6.15%	-0.24	-16.61%

2	Growth Stocks	7.08%	14.41%	0.49	-23.26%
3	Growth + Value Stocks	3.97%	8.60%	0.46	-13.42%
4	Overall	3.13%	6.61%	0.47	-11.61%

No.	Investment Style	Total Return (%)	Avg. Monthly Return (%)	Avg. Quarterly Return (%)
1	Value Stocks	-7.52%	-0.18%	-0.50%
2	Growth Stocks	31.32%	0.82%	2.36%
3	Growth + Value Stocks	17.97%	0.44%	1.34%
4	Overall	14.40%	0.37%	1.11%



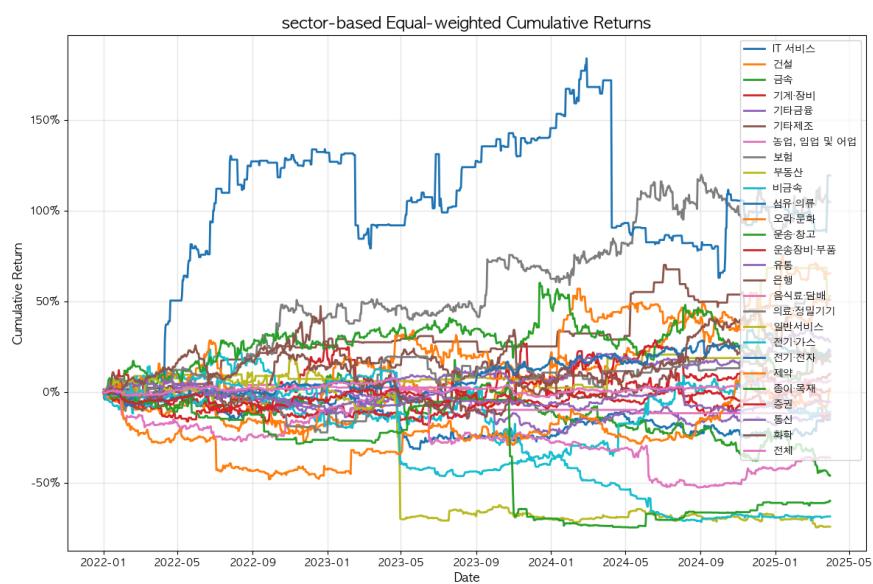


Performance by Industry Sector (Based on KRX Classification)

Note: Industry classification is based on official sector definitions provided by the Korea Exchange (KRX), instead of the standard GICS framework.

No.	Sector	Total Trades	Avg. Daily Traded Stocks	Annualized Return	Annualized Volatility	Sharpe Ratio	Max Drawdown
1	Electrical & Electronics	26,219	22.13	4.91%	9.49%	0.52	-11.1%
2	Retail	19,488	16.45	5.92%	10.49%	0.56	-11.0%
3	Securities	18,417	15.54	0.57%	8.80%	0.06	-10.1%
4	Other Financials	15,697	13.25	-2.53%	10.96%	-0.23	-25.1%
5	Transport Equipment & Parts	15,683	13.23	2.88%	12.61%	0.23	-24.1%

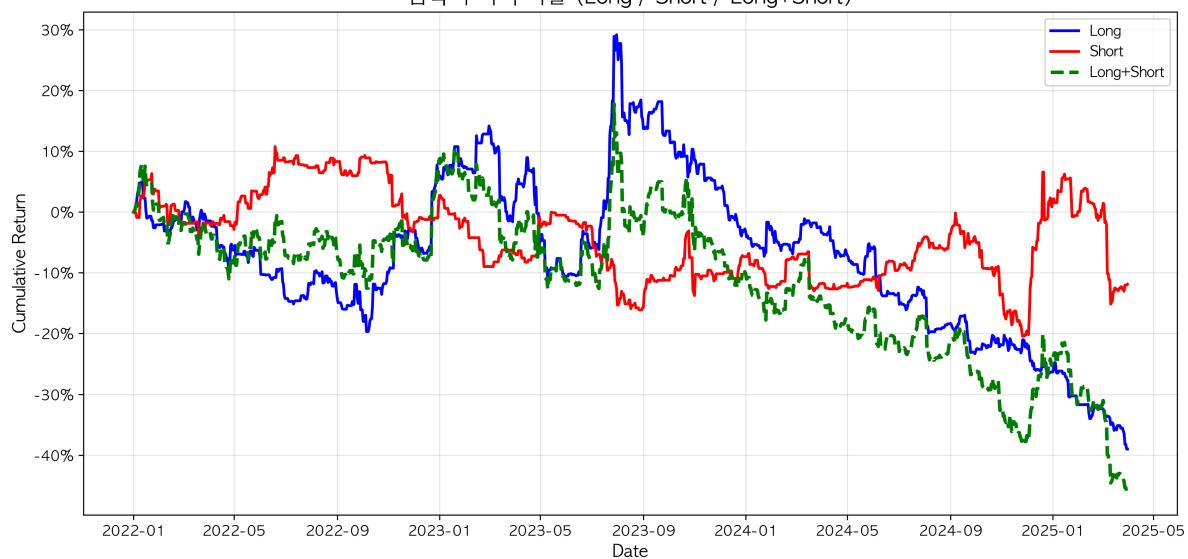
6	Chemicals	10,673	9.01	5.45%	15.54%	0.35	-32
7	IT Services	7,986	6.74	-2.23%	15.54%	-0.14	-34
8	Food & Beverage	7,524	6.35	-6.96%	20.41%	-0.34	-54
9	Construction	6,207	5.24	10.89%	15.99%	0.68	-24.
10	Metals	5,463	4.61	-10.17%	21.76%	-0.47	-54
11	Transport & Warehousing	4,328	3.65	5.77%	19.18%	0.30	-34
12	Insurance	4,291	3.62	18.06%	16.48%	1.10	-15.
13	Pharmaceuticals	2,896	2.44	3.52%	24.62%	0.14	-50
14	General Services	2,808	2.37	-19.59%	33.73%	-0.58	-78.
15	Machinery & Equipment	2,490	2.10	0.59%	26.42%	0.02	-34
16	Non-metallic Minerals	2,129	1.80	1.00%	26.59%	0.04	-48
17	Telecommunications	—	1.58	-0.69%	9.63%	-0.07	-17.8
18	Entertainment & Culture	1,859	1.57	14.27%	23.00%	0.62	-43
19	Banking	1,844	1.56	5.78%	18.54%	0.31	-34
20	Utilities (Electric & Gas)	1,695	1.43	-19.49%	23.55%	-0.83	-74.
21	Real Estate	782	0.66	10.77%	16.26%	0.66	-10.
22	Textiles & Apparel	646	0.55	22.72%	26.60%	0.85	-42
23	Paper & Forest Products	416	0.35	-14.17%	27.99%	-0.51	-74.
24	Medical & Precision Devices	300	0.25	5.19%	18.92%	0.27	-31.
25	Other Manufacturing	187	0.16	9.64%	9.60%	1.00	-15.
26	Agriculture, Forestry & Fishery	22	0.02	-2.48%	4.47%	-0.55	-14.

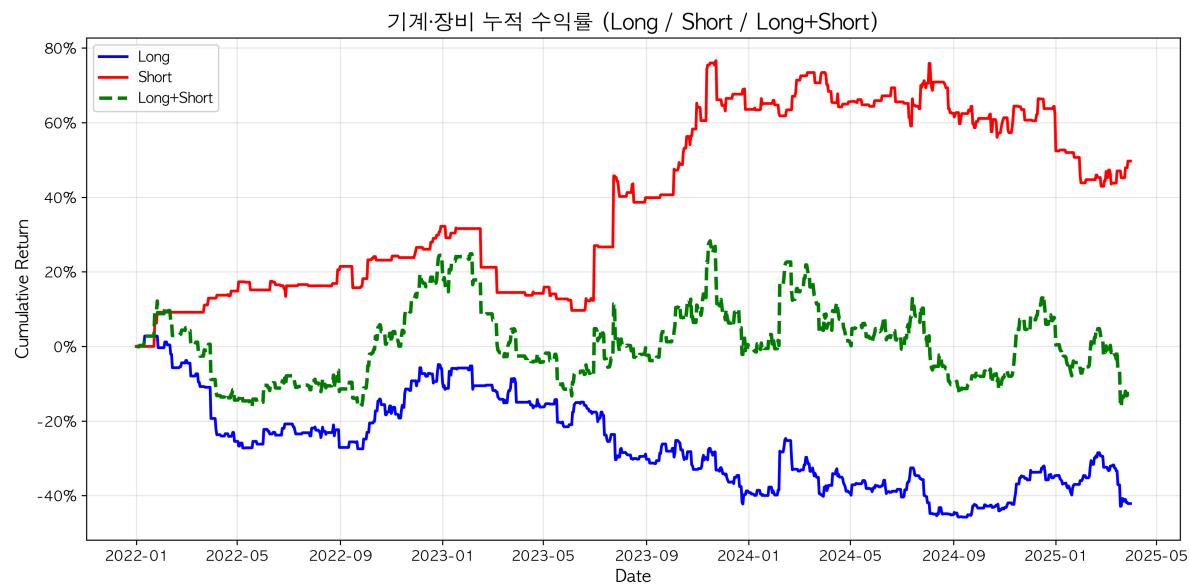


건설 누적 수익률 (Long / Short / Long+Short)

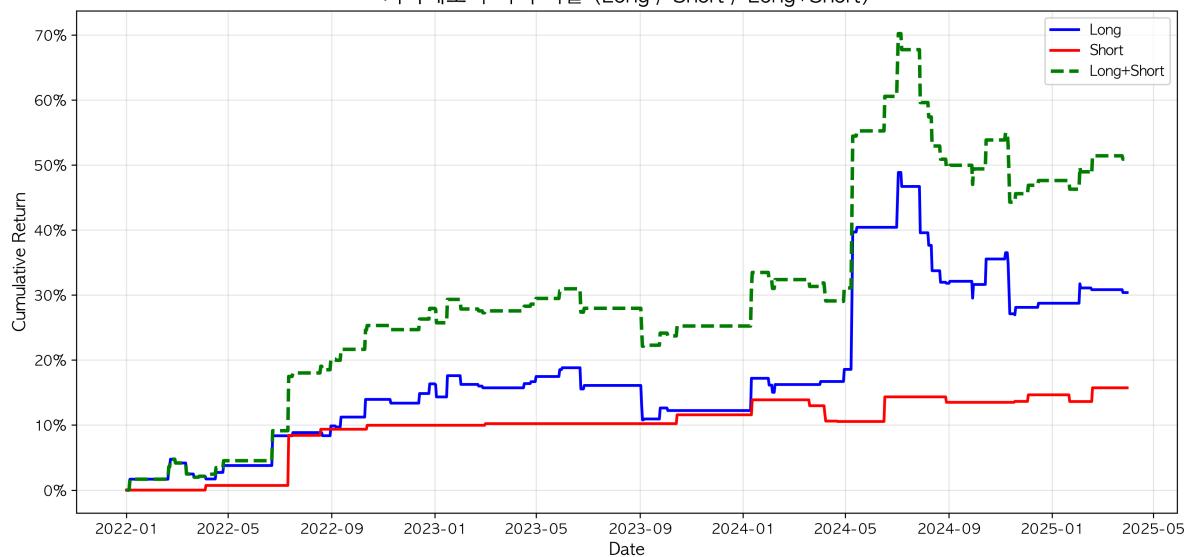


금속 누적 수익률 (Long / Short / Long+Short)

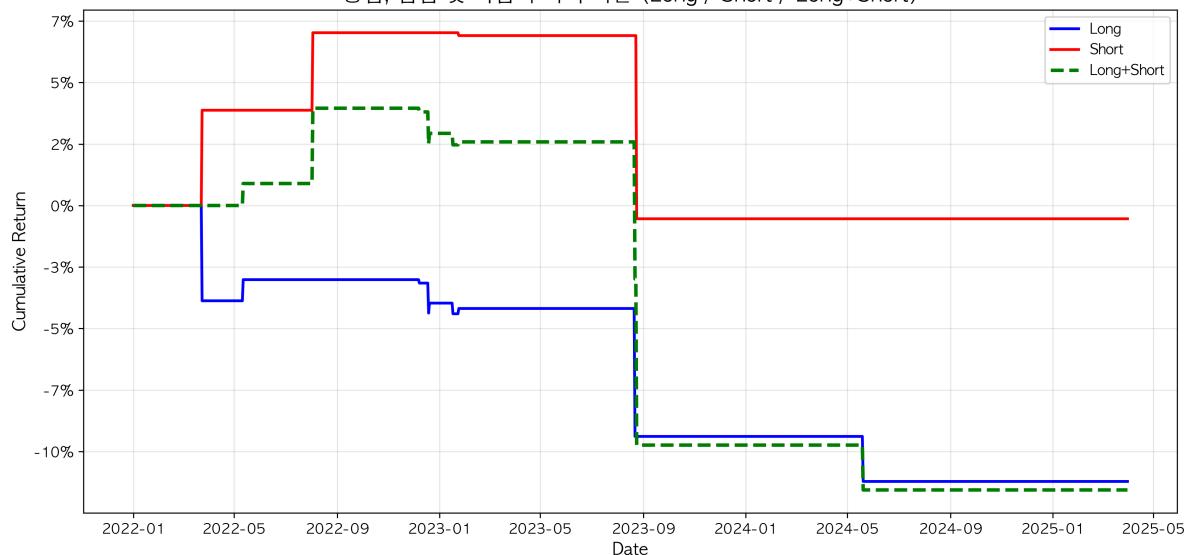




기타제조 누적 수익률 (Long / Short / Long+Short)



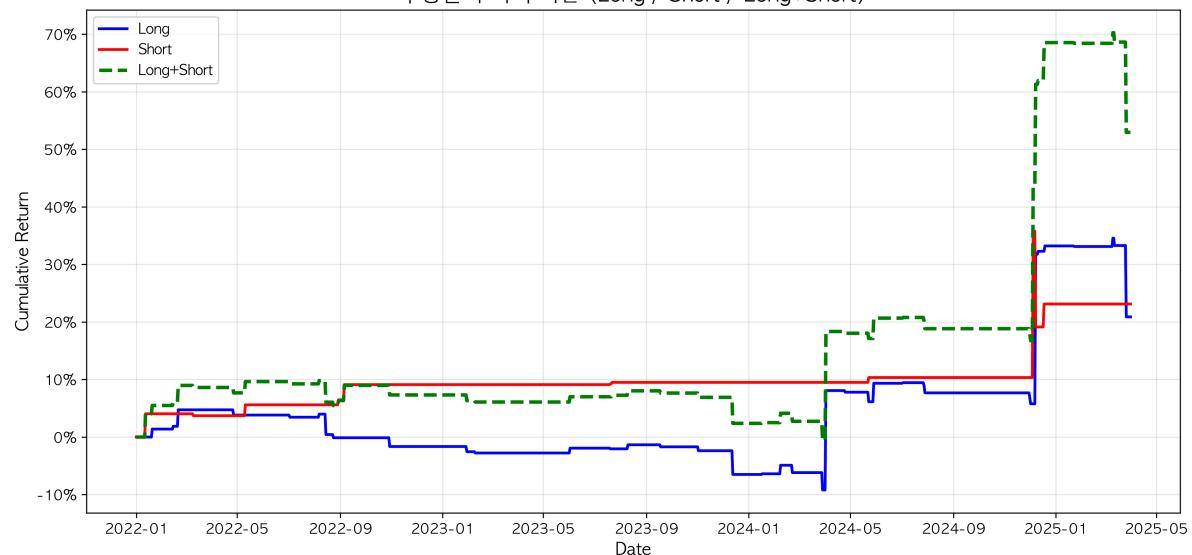
농업, 임업 및 어업 누적 수익률 (Long / Short / Long+Short)



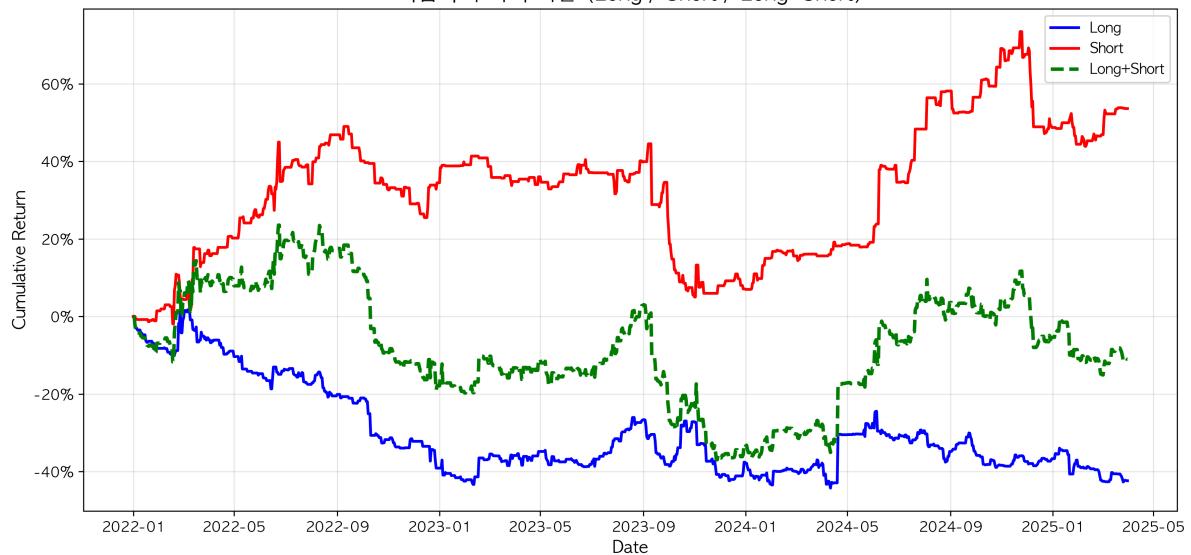
보험 누적 수익률 (Long / Short / Long+Short)



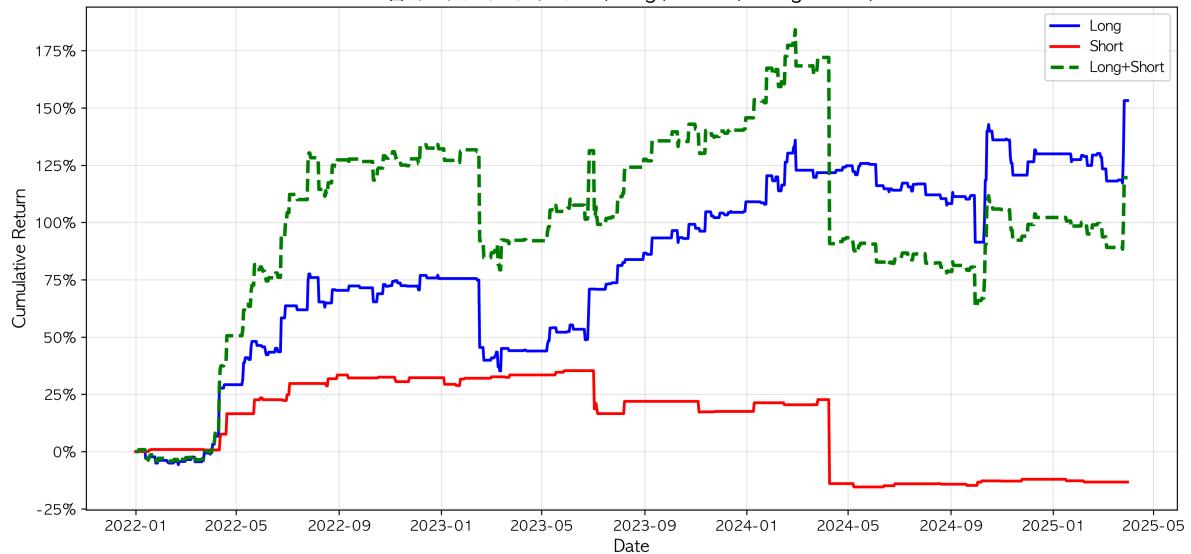
부동산 누적 수익률 (Long / Short / Long+Short)

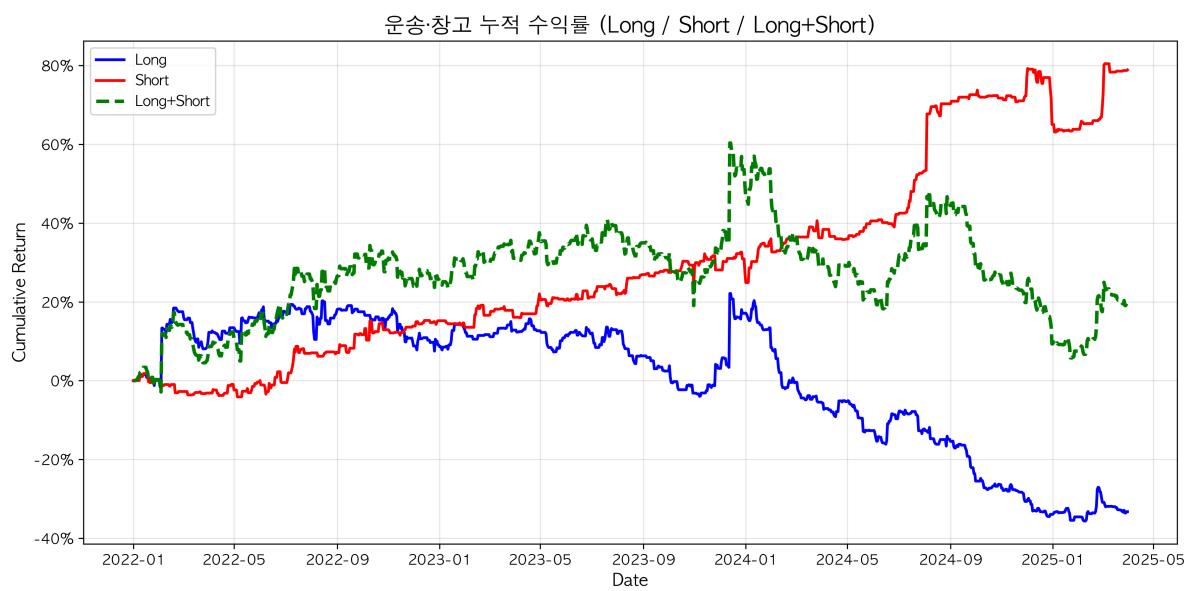
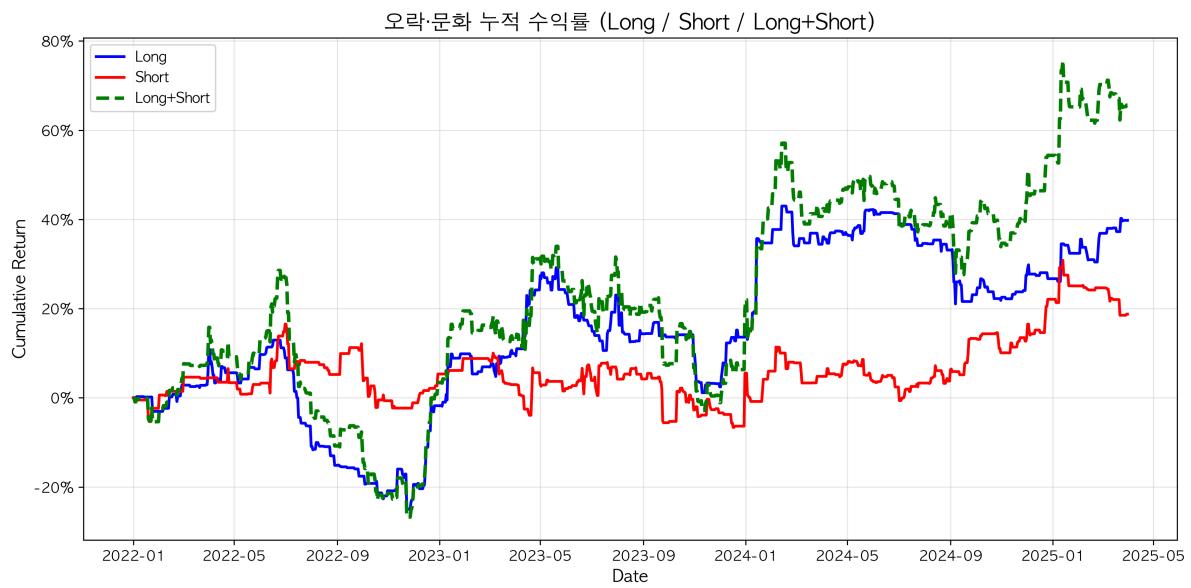


비금속 누적 수익률 (Long / Short / Long+Short)

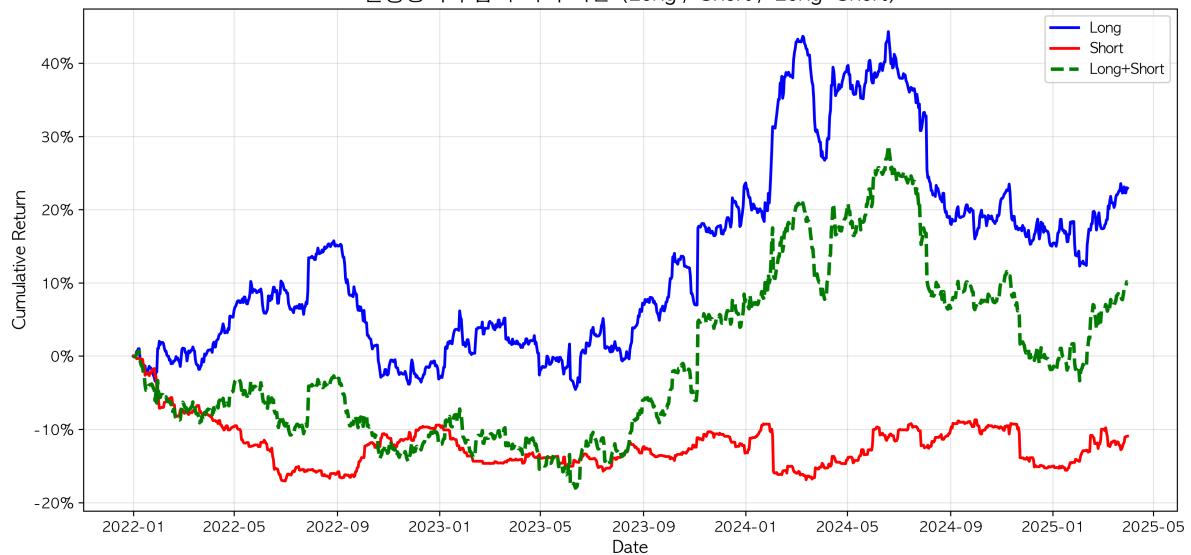


섬유·의류 누적 수익률 (Long / Short / Long+Short)



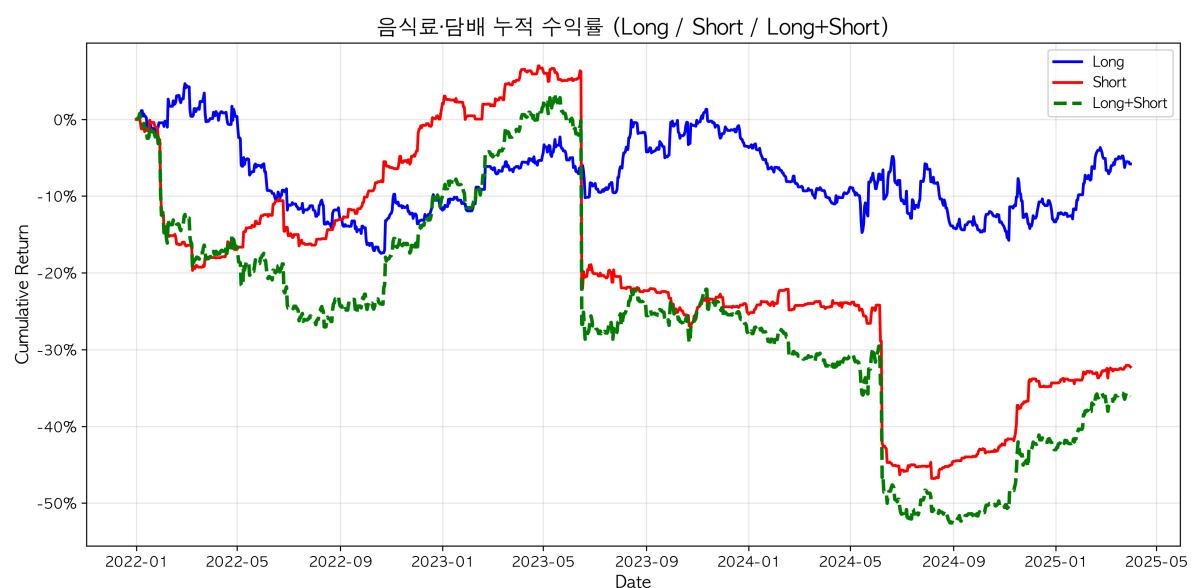
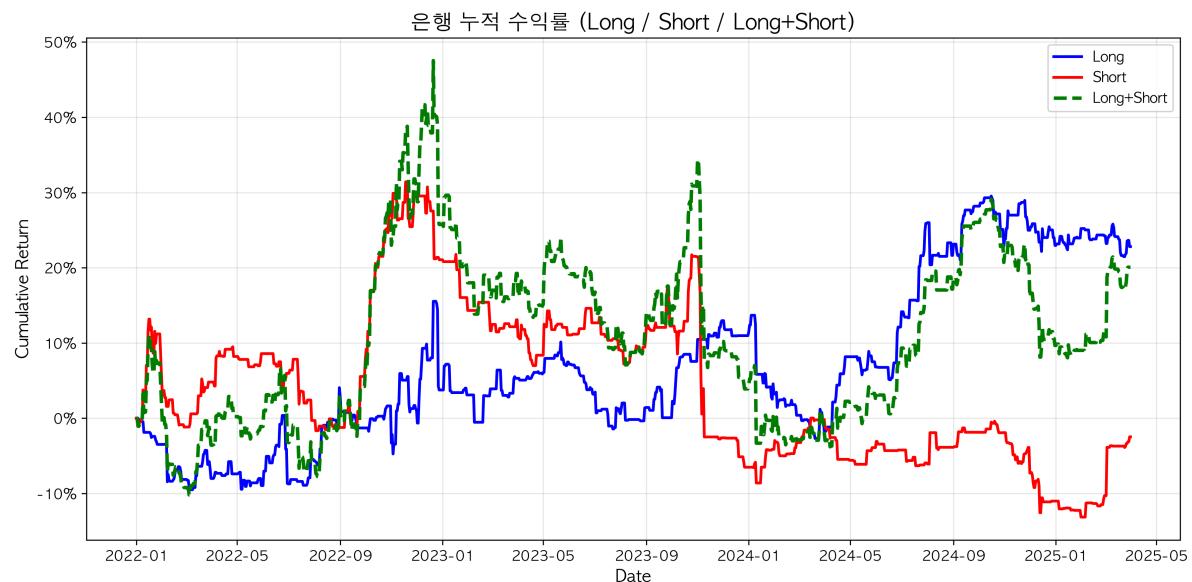


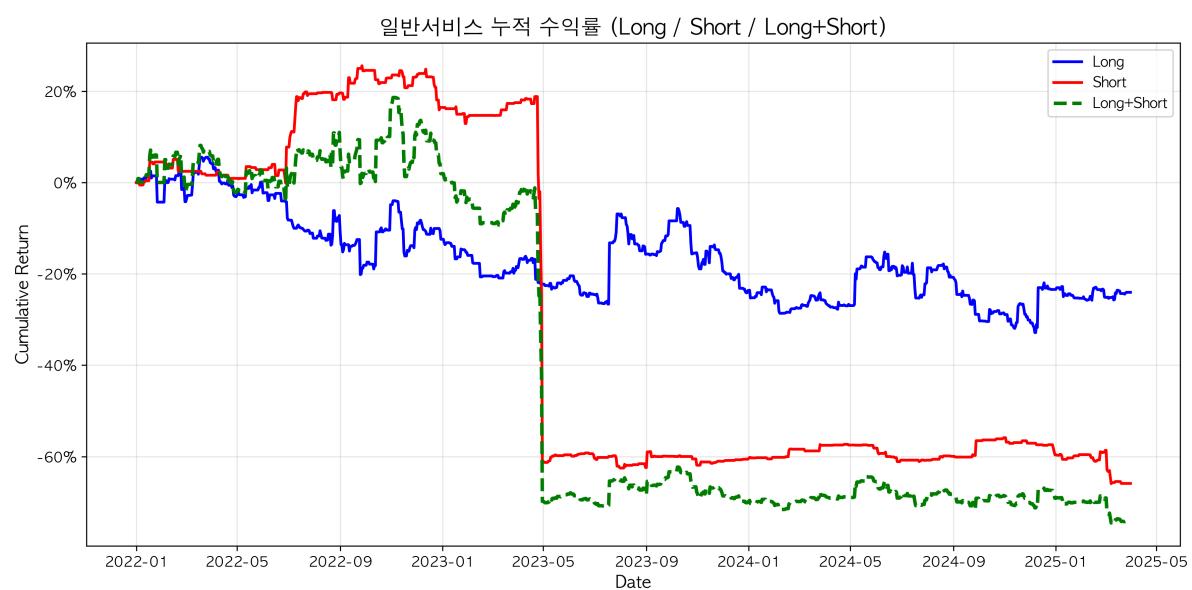
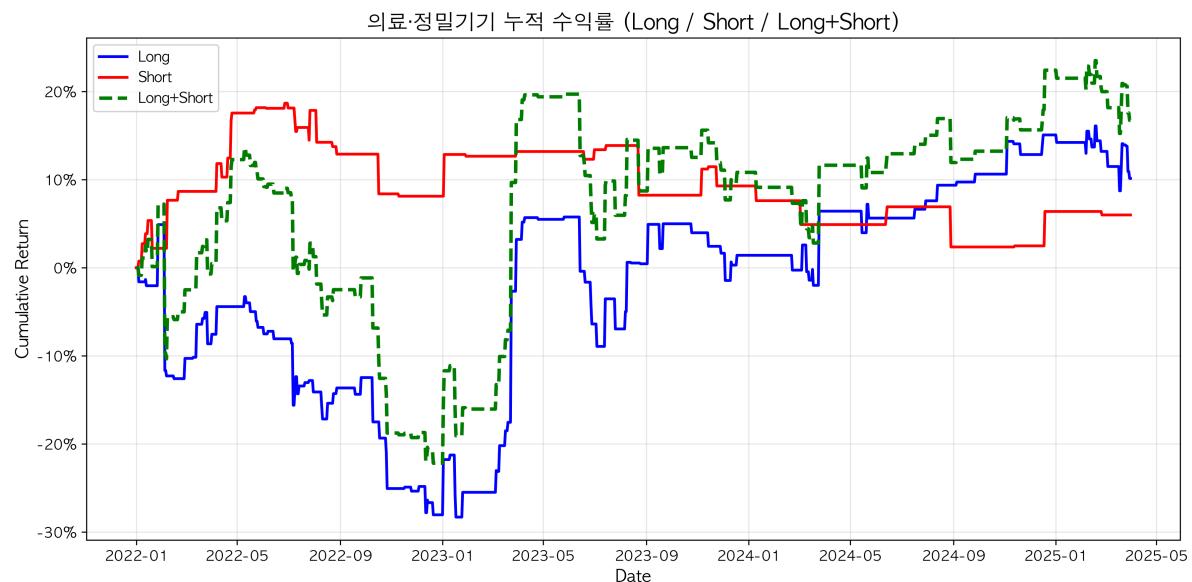
운송장비·부품 누적 수익률 (Long / Short / Long+Short)

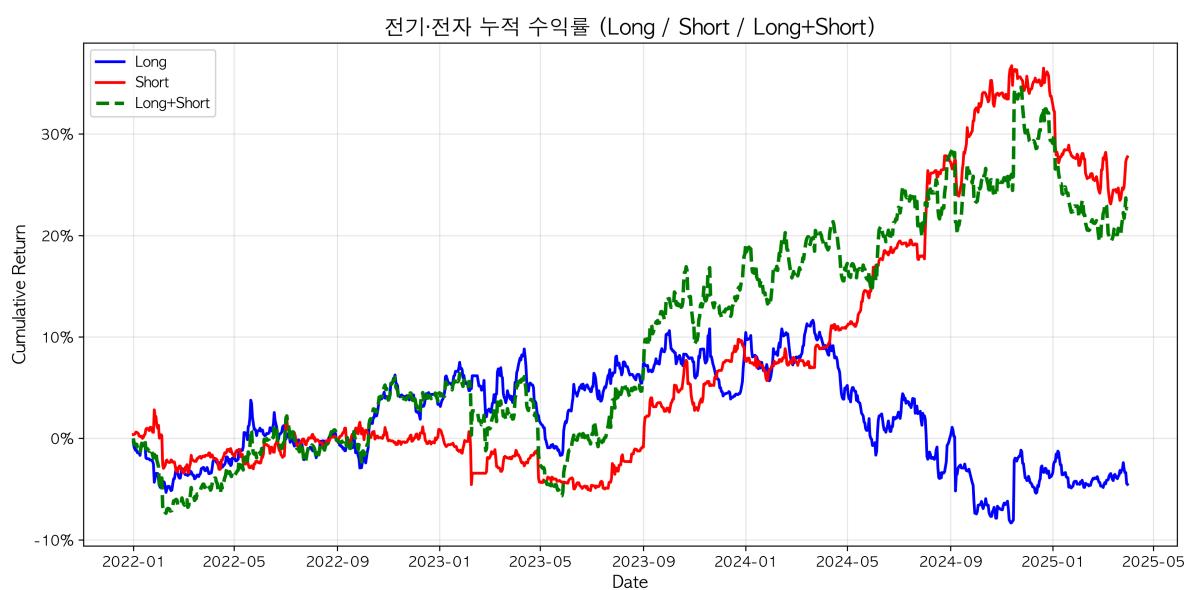
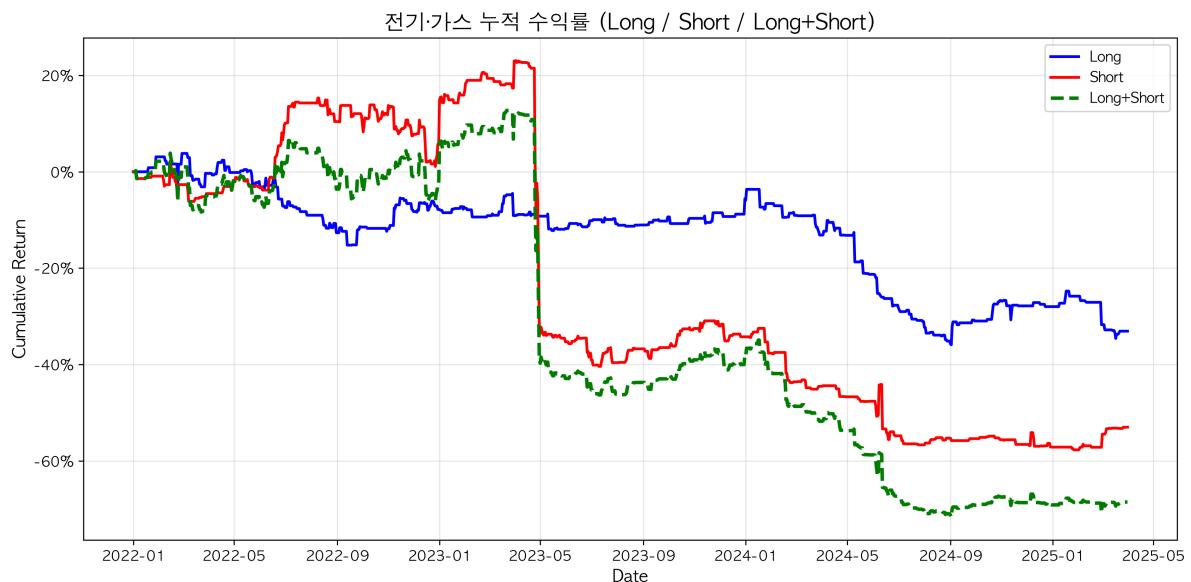


유통 누적 수익률 (Long / Short / Long+Short)





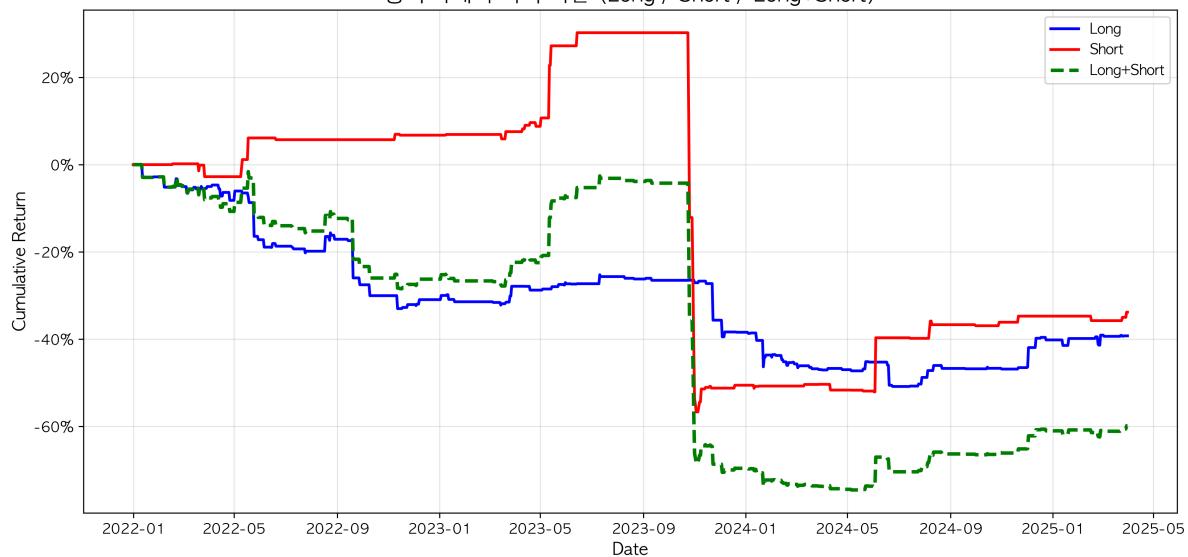




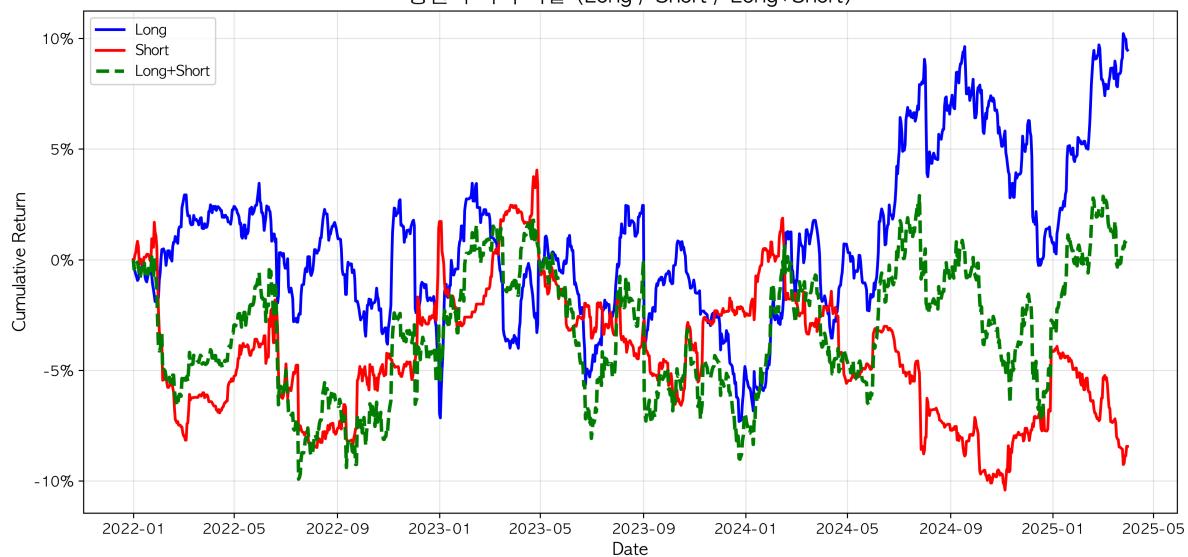
제약 누적 수익률 (Long / Short / Long+Short)



종이·목재 누적 수익률 (Long / Short / Long+Short)

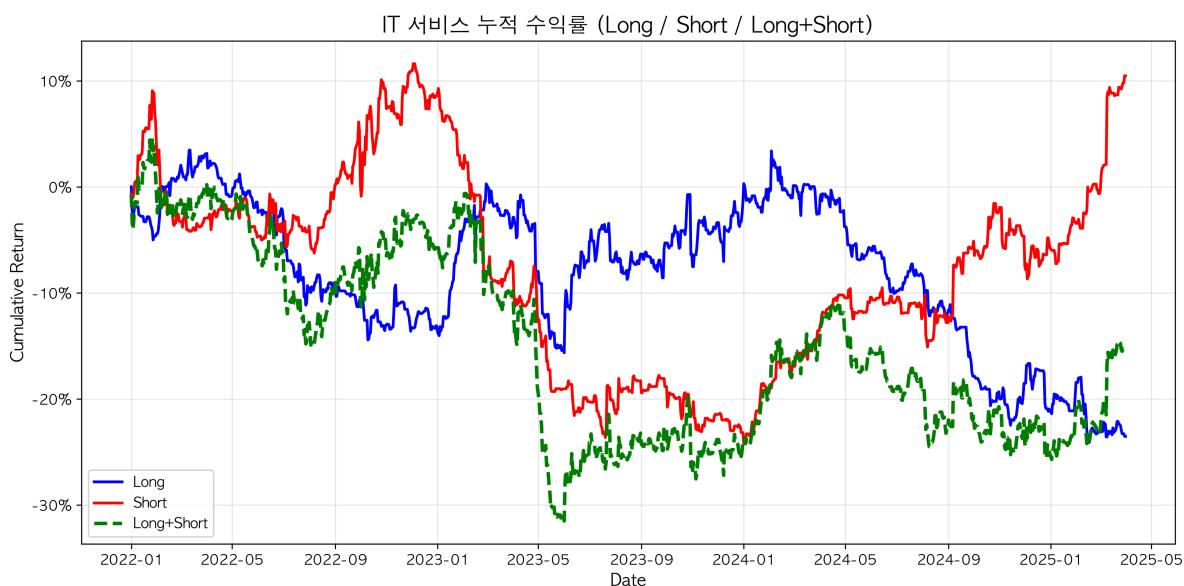
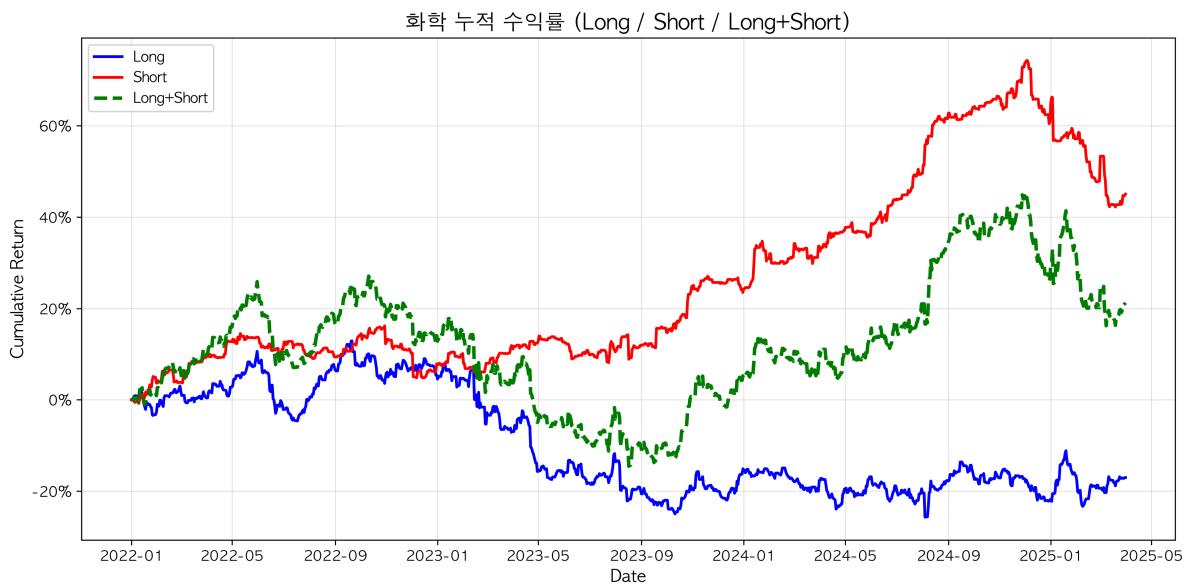


증권 누적 수익률 (Long / Short / Long+Short)



통신 누적 수익률 (Long / Short / Long+Short)





Conclusion

Conclusion: Evaluation of a GPT-Based Sentiment-Driven Investment Strategy

1 Model and Strategy Overview

This strategy utilizes **GPT-4.1 Nano** to extract sentiment scores from news headlines, enabling a **systematic long-short trading strategy** driven by daily sentiment signals.

Unlike earlier GPT-3 models that emphasized surface-level sentiment, GPT-4.1 Nano provides **deeper contextual understanding and neutral interpretations**.

2 Overall Strategy Performance

Standalone Strategy Results (Long / Short / Long-Short)

- Long-only annualized return: +1.79%
- Short-only annualized return: +2.35%
- Long-short combined return: -0.54%
 - Due to **loss-weighted exposure**, where losing sides dominated on specific days.

Correlation Analysis

- GPT score vs Daily return: +0.015
- Headline length vs Return: 0.002
- ➤ These **very weak correlations** indicate that **direct sentiment signals are insufficient predictors of returns**, especially when generated by a more nuanced language model.

3 Performance by Segment

By Sector (KRX Classification)

- Top performers:
 - **Textiles & Apparel** (+119.55%)
 - **Insurance** (+104.97%, Sharpe 1.10)
 - **Construction, Retail, Electrical & Electronics, Entertainment** also posted strong results.
- Underperformers:
 - **General Services, Food & Beverage, Metals, Utilities, Forest Products**
 - Characterized by **high volatility** and **deep drawdowns** (up to -78%).

By News Category

- **General Economy** category led with:
 - **15.59% annualized return, Sharpe 0.91**
- Categories like **Resources, Automotive, Real Estate, Semiconductors** performed poorly.

By Investment Style

- **Growth stocks** outperformed:
 - **+7.08% annualized return, 31.32% total return**
- **Value stocks** underperformed with a **-7.52% total return**

By Firm Size

- Only **large-cap** stocks delivered positive returns:
 - **2.62% annualized return**, with moderate drawdowns (-11.16%)
- **Small- and mid-cap** stocks saw negative returns and higher volatility.

4 Key Takeaways & Strategy Implications

Insight	Explanation
1. Better language models ≠ better predictive power	GPT-4.1 Nano's nuanced interpretations reduce the direct link between sentiment polarity and price movement.
2. Industry selection is critical	Strategy performance varies drastically across sectors . Sector filtering is essential.
3. Growth & large-cap bias is favorable	Strategies focusing on growth-oriented and large-cap stocks show better risk-adjusted returns.
4. Portfolio construction needs improvement	Equal or unbalanced exposure to long/short sides causes downside. Dynamic weight adjustment is needed.
5. Sentiment signal needs sophistication	Rather than simple thresholds, inference-based triggers, reinforcement learning, or context-aware sentiment modeling may enhance predictive power.

