

---

# **Datago- Retail Investor Sentiment Analytics- South Korea (RISA-Korea)**

Last updated on 13 Feb 2026

---

[info@datago.com.hk](mailto:info@datago.com.hk)

# Background- South Korea Retail Investors

- Almost one-fifth of Korea's population of 52 million dabbles in stocks, and data showed local brokerages have amassed Won76tn in cash deposits.
- 64% of annual transaction amounts in the Korean stock markets were accounted by retail investors in 2022.
- The number of retail investors reached over 14.24 million in 2022, more than double the number of retail investors prior to the outbreak of COVID-19 in 2020.



# Background- Naver

Naver was launched in 1999 as the first web portal in South Korea to develop and use its own search engine. In addition to offering a search results, Naver also integrates different features, including news and forums.

As of September 2017, Naver handled 74.7% of all web searches in South Korea and had 42 million enrolled users. More than 25 million Koreans have Naver as the start page on their default browser and the mobile application has 28 million daily visitors.

Naver's stock forum has now recorded over 106 million posts and 80 million replies since 2017.

The screenshot shows a Naver stock forum post for Samsung Electronics (삼성전자 005930). The stock price is 79,900 KRW, up 1.65% from the previous day's closing price of 78,600 KRW. The post title is '종가 8만전자 80,300원' (Closing price 80,300 KRW for 80,000 shares). The post content includes a prediction: '? 예상치' (Prediction: ?). The post has 2 likes and 0 replies. A comment section is visible with a message: '댓글을 작성하려면 로그인해주세요' (Please log in to write a comment). A notice at the bottom states: '클린봇이 악성댓글을 감지합니다.' (Cleanbot detected malicious comments).

삼성전자 005930 코스피	2024.06.14 15:23 기준(장중)	일시간	기업개요
79,900	전일 78,600   고가 80,500 (상한가 102,100)   거래량 21,240,076		
전일대비 ▲1,300   +1.65%	시가 79,700   저가 79,000 (하한가 55,100)   거래대금 1,696,026 백만		

종합정보 | 시세 | 차트 | 투자자별 매매동향 | 뉴스공시 | 종목분석 | **종목토론실** | 전자공시 | 공모도현황

종목토론실 | 삼성전자 주주오픈톡 > | ① | 답변하기 | 목록

종가 8만전자 80,300원 | 조회 67 | 공감 2 | 비공감 0 | 신고

dlxo\*\*\*\* 223.39.\*\*\*.202 | 작성자를 더보기 > | 2024.06.14 15:18

? 예상치

2 | 0

댓글 1 | 🔄

댓글을 작성하려면 로그인해주세요

클린봇이 악성댓글을 감지합니다. | ⚙️ 설정

# Background- TossInvest

TossInvest was launched in early 2021 as a mobile-first investment platform. In addition to trading and portfolio management functions, TossInvest integrates an in-app investor community that allows users to exchange opinions and information.

Leveraging the large user base of the Toss super-app—widely used in South Korea for payments, banking, and personal finance—the TossInvest forum has grown rapidly. The platform is particularly popular among younger retail investors, contributing to a high level of real-time interaction within the forum.

TossInvest forum has now recorded over 8 million posts and 4 million replies since 2021.



# Background-TossInvest

## Features:

- Enable direct stock trading through Toss Securities
- With user consent, selected real trading records will be shared publicly as posts
- Posts based on actual trades will include key transaction details
  - ✓ Execution time
  - ✓ Trade price
  - ✓ Quantity traded
  - ✓ Profit and loss (P&L)
  - ✓ Return rate

The image shows two versions of a stock trading interface side-by-side. The left version is in Korean, and the right version is in English. Both interfaces have the following fields and options:

- 주문하기 / Place Order**
- 주문 유형 / Order Type:** 일반 주문 / Normal Order
- 구매 가격 / Buy Price:** 지정가 (selected) / 시장가 / Limit (selected) / Market
- 수량 / Quantity:** 수량 입력 / Enter quantity
- Price/Quantity Selection:** 93,500 원 / 93,500 KRW, with minus and plus buttons.
- Additional Options:** 10%, 25%, 50%, 최대 / 10%, 25%, 50%, Max
- Summary:** 구매가능 금액 0원 / Purchasable Amount 0 KRW, 총 주문 금액 0원 / Total Order Amount 0 KRW
- Action Button:** 로그인하고 구매하기 / Log In and Purchase



# Key Facts of RISA-Korea

\*As of 2025-12-31

	#Users	#Posts	#Covered Security	#Stocks with over 300 posts/month (avg. 10 posts/day)	#Rebalances	#Stocks with over 30 rebalances/month (avg. 1 rebalances/day)
Full History since 2017 (Naver)	2.1M	106M	5016	N/A		N/A
1 month (Dec 2025) (Naver)	105K	1.3M	3562	653		
Full History since 2021 (TossInvest)	475K	8.5M	4000	N/A	824K	N/A
1 month (Dec 2025) (TossInvest)	64K	457K	2848	233	33K	193

## Coverage:

5000+ Korea securities  
(3100+stocks, 1600+ ETP)

## Update Frequency:

Daily (T+1, 02:00, UTC+9)

## History:

- Naver: 2017-06-07
- TossInvest: 2021-06-09

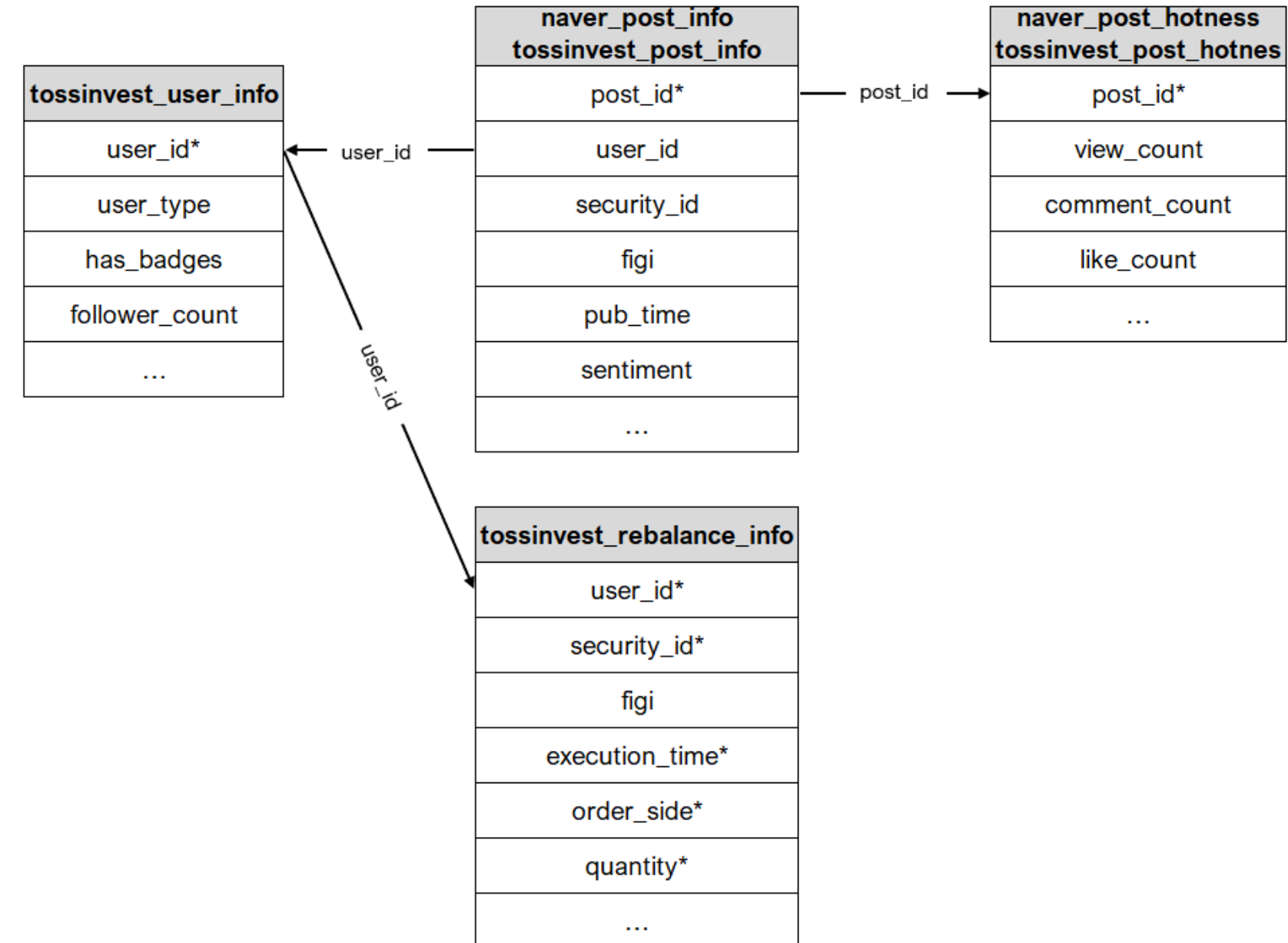
## Realtime PIT:

- Naver: 2024-05-01
- TossInvest: 2026-01-01

# Data Structure of RISA-Korea

naver_stats_basic tossinvest_stats_post
security_id*
figi
pub_date*
view_neg_sum
comment_neg_sum
like_neg_sum
user_neg_sum
senti_neg_sum
senti_conform
...

tossinvest_stats_rebalance
security_id*
figi
pub_date*
buy_user_num
buy_rebalance_num
buy_rebalance_share
buy_rebalance_price_avg
buy_rebalance_price_min
buy_rebalance_price_max
sell_user_num
...



3 table of aggregated analytics in stock level

6 tables of fine-grained information in record level

# Data Tables-Aggregate Statistics (Selected Fields)

## 1 RISA-Korea-V3-naver\_stats\_basic

Field Name	Description	Sample
security_id	Security code in Korea	000020
figi	FIGI code of the security	BBG000BDH722
pub_date	publish date, 2020-07-05 means data from 2020-07-05 00:00:00 to 2020-07-05 23:59:59, can be accessed after 2020-07-06 02:00:00, in UTC+9.	2023-11-01
view_neg_sum	The total number of views for all negative posts.	3326
senti_neg_sum	The sum of sentiment score of negative posts.	-24.090959
senti_score_div	Sentiment ratio index: (number of positive messages - number of negative messages)/ (number of all messages). The values ranges from -1 to 1, and higher value indicates stronger positive sentiment by measuring the percentage of positive messages among all messages	-0.833333
senti_score_log	Logarithm sentiment index: $\ln[(1 + \text{number of positive messages}) / (1 + \text{number of negative messages})]$ . Higher value indicates stronger positive sentiment by measuring (logarithm) advantage of positive messages over negative messages	-1.79176
senti_conform	Sentiment conformity index: $1 - (1 - \text{sentiment ratio index}^2)^{0.5}$ . The value ranges from 0 to 1, higher value indicates more users agree with each other on the sentiment (either positive or negative)	0.447229
...	...	...

# Data Tables-Aggregate Statistics (Selected Fields)

2

## RISA-Korea-V3-tossinvest\_stats\_post

Field Name	Description	Sample
security_id	Security code in Korea	000250
figi	FIGI code of the security	BBG000CBJTV3
pub_date	publish date, 2020-07-05 means data from 2020-07-05 00:00:00 to 2020-07-05 23:59:59, can be accessed after 2020-07-06 02:00:00, in UTC+9.	2025-06-30
view_neg_sum	The total number of views for all negative posts.	1331
senti_neg_sum	The sum of sentiment score of negative posts.	-9.45304
senti_score_div	Sentiment ratio index: (number of positive messages - number of negative messages)/ (number of all messages). The values ranges from -1 to 1, and higher value indicates stronger positive sentiment by measuring the percentage of positive messages among all messages	-0.174957
senti_score_log	Logarithm sentiment index: $\ln[(1 + \text{number of positive messages}) / (1 + \text{number of negative messages})]$ . Higher value indicates stronger positive sentiment by measuring (logarithm) advantage of positive messages over negative messages	-1.50408
senti_conform	Sentiment conformity index: $1 - (1 - \text{sentiment ratio index}^2)^{0.5}$ . The value ranges from 0 to 1, higher value indicates more users agree with each other on the sentiment (either positive or negative)	0.015424
...	...	...

---

# Data Tables-Aggregate Statistics (Selected Fields)

3

## RISA-Korea-V3-tossinvest\_stats\_rebalance

Field Name	Description	Sample
security_id	Security code in Korea	000720
figi	FIGI code of the security	BBG000BCVC63
pub_date	publish date, 2020-07-05 means data from 2020-07-05 00:00:00 to 2020-07-05 23:59:59, can be accessed after 2020-07-06 02:00:00, in UTC+9.	2025-06-30
buy_user_num	The number of users that buy the stock.	5
buy_rebalance_num	The number of rebalance actions that buy the stock.	8
buy_rebalance_share	Total shares of rebalance actions that buy the stock.	280
buy_rebalance_price_avg	Average price of rebalance actions that buy the stock.	78562.5
buy_rebalance_price_min	Minimum price of rebalance actions that buy the stock.	76700
buy_rebalance_price_max	Maximum price of rebalance actions that buy the stock.	81100
buy_rebalance_value	Total market value of rebalance actions that buy the stock.	22486600
sell_user_num	The number of users that sell the stock.	4
...	...	...

---

# Data Tables-Underlying Information (Selected Fields)

## 4 RISA-Korea-V3-naver\_post\_info

Field Name	Description	Sample
post_id	Hashed post id	e8b0c4c7e128a9a7
user_id	Hashed user id	5f9266a512f2f608
security_id	Security code in Korea	019660
pub_time	Publish time of post in UTC+9 as recorded in the platform	2019-01-01 00:00:07
sentiment	Sentiment score of posts based on Datago's text analysis of post content. The sentiment score ranges from -1 to 1, where -1 represents the most negative sentiment and 1 represents the most positive sentiment.	0.191527

## 5 RISA-Korea-V3-naver\_post\_hotness

Field Name	Description	Sample
post_id	Hashed post id	02941f549070a982
view_count	The number of view of the post	965
comment_count	The number of reply of the post	0
like_count	The number of like of the post	2
dislike_count	The number of dislike of the post	3

# Data Tables-Underlying Information (Selected Fields)

6

## RISA-Korea-V3-tossinvest\_post\_info

Field Name	Description	Sample
post_id	Hashed post id	664b75bfb7c666eb
user_id	Hashed user id	4072140f1110c5c7
security_id	Security code in Korea	475560
pub_time	Publish time of post in UTC+9 as recorded in the platform	2025-05-31 00:06:34
holding_flag	Indicates whether the user is currently holding the stock, disclosed by tossinvest; 1 means yes, and 0 means no	1
sentiment	Sentiment score of posts based on Datago's text analysis of post content. The sentiment score ranges from -1 to 1, where -1 represents the most negative sentiment and 1 represents the most positive sentiment.	-0.172053

7

## RISA-Korea-V3-tossinvest\_post\_hotness

Field Name	Description	Sample
post_id	Hashed post id	d657c83fa7d541e6
view_count	The number of view of the post	608
comment_count	The number of reply of the post	3
like_count	The number of like of the post	10
repost_count	The latest repost count of the post available before the delivery_time	2

---

# Data Tables-Underlying Information (Selected Fields)

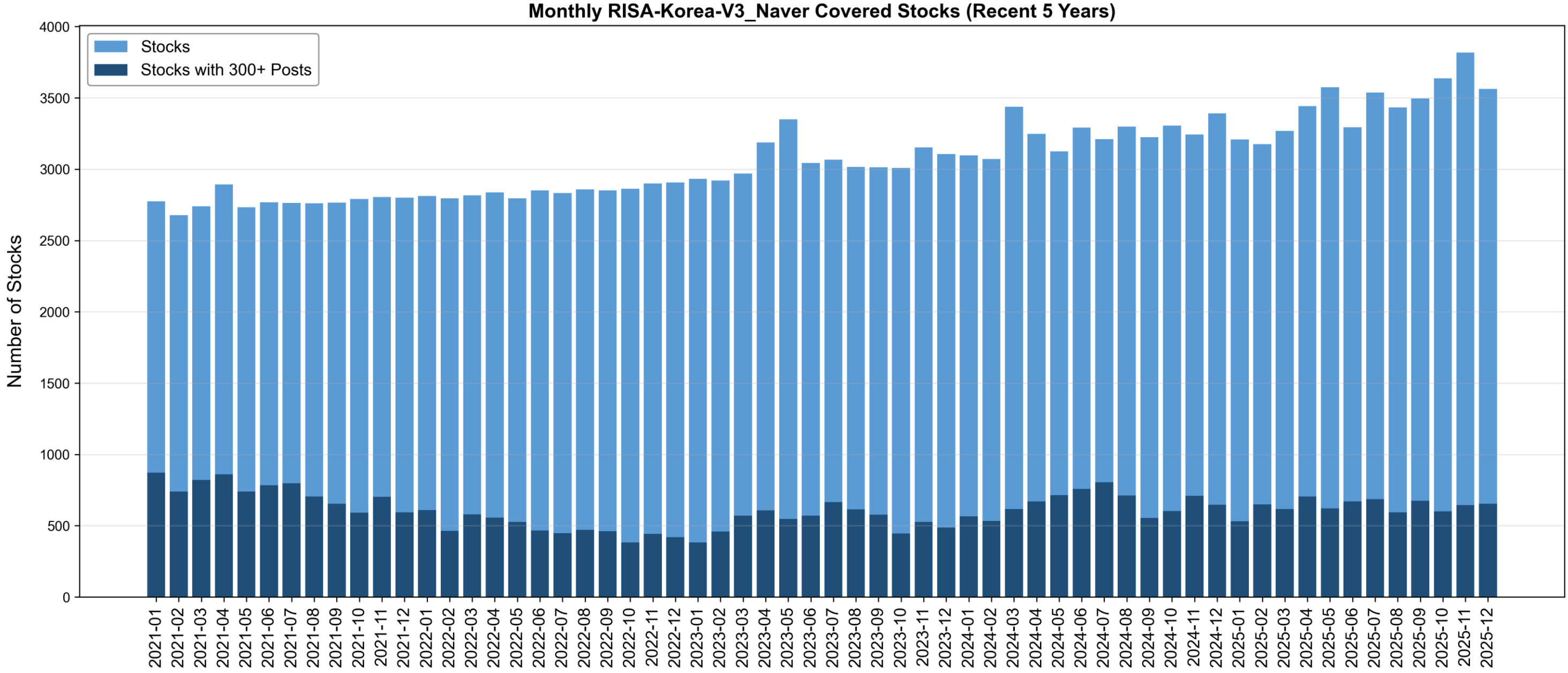
## 8 RISA-Korea-V3-tossinvest\_rebalance\_info

Field Name	Description	Sample
user_id	Hashed user id	ee3907599042407a
security_id	Security code in Korea	103590
execution_time	Execution time of the rebalance	2025-06-30 08:02:13
order_side	The direction of the order, indicating whether it is a buy or sell transaction. 0 means sell and 1 means buy.	0
quantity	The number of units of the security involved in the transaction	450
average_price	The average price in KRW at which the securities were bought or sold	38800
profit_amount	The total profit in KRW realized from the transaction	3942026.0

## 9 RISA-Korea-V3-tossinvest\_user\_info

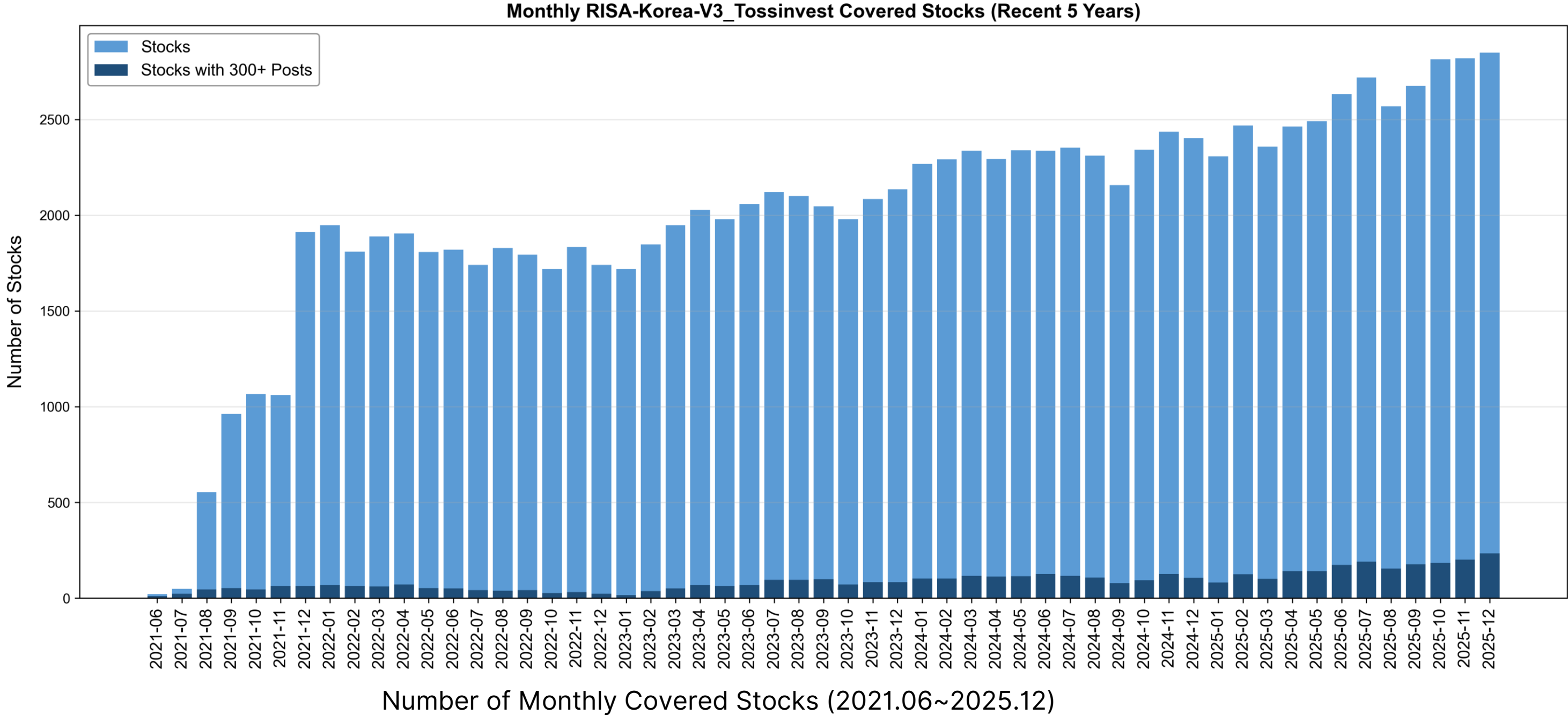
Field Name	Description	Sample
user_id	Hashed user id	8dac4c8cfa1e216e
user_type	User type; 1 means normal user, 2 means partner, 3 means companion	1
has_badges	Indicates whether the user has badges or not; 1 means yes, and 0 means no	0
follower_count	Follower count of the user	46

# Trend of Covered Stocks -Naver

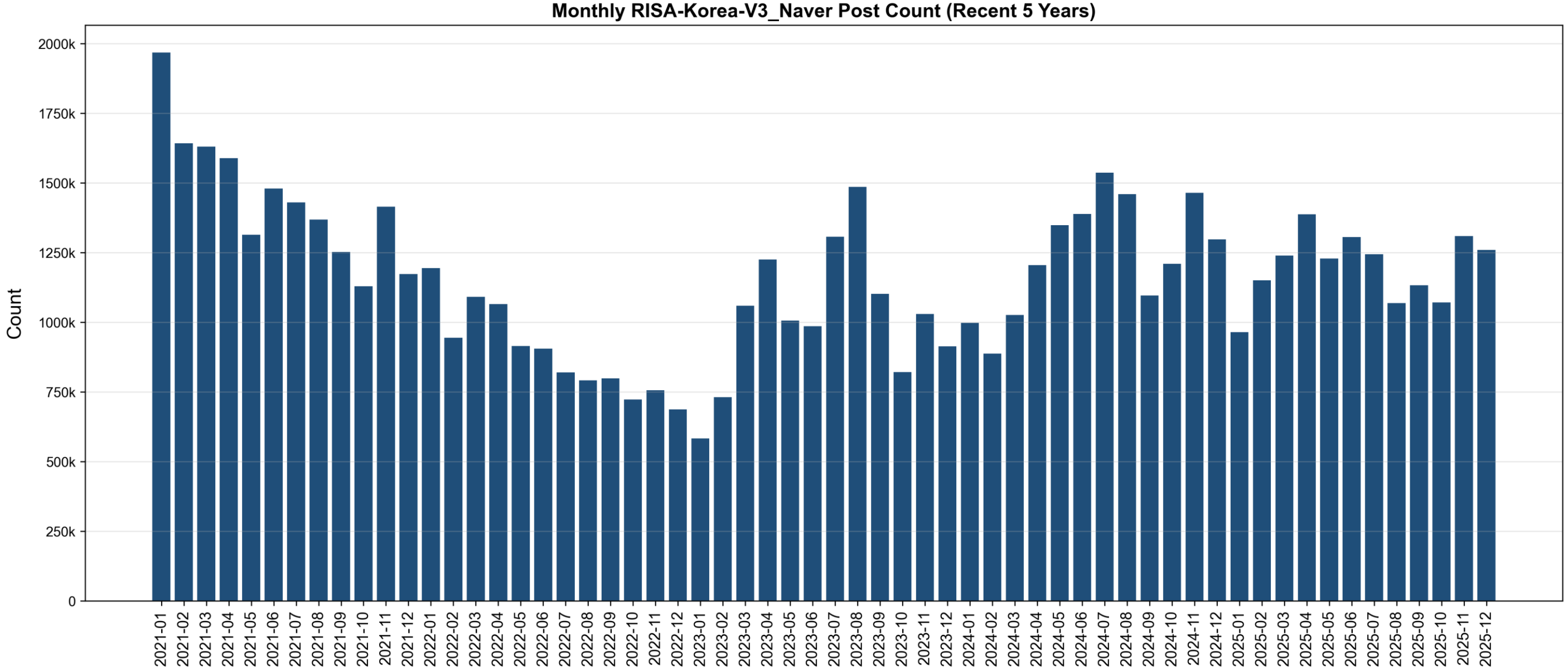


Number of Monthly Covered Stocks (2021.01~2025.12)

# Trend of Covered Stocks -TossInvest

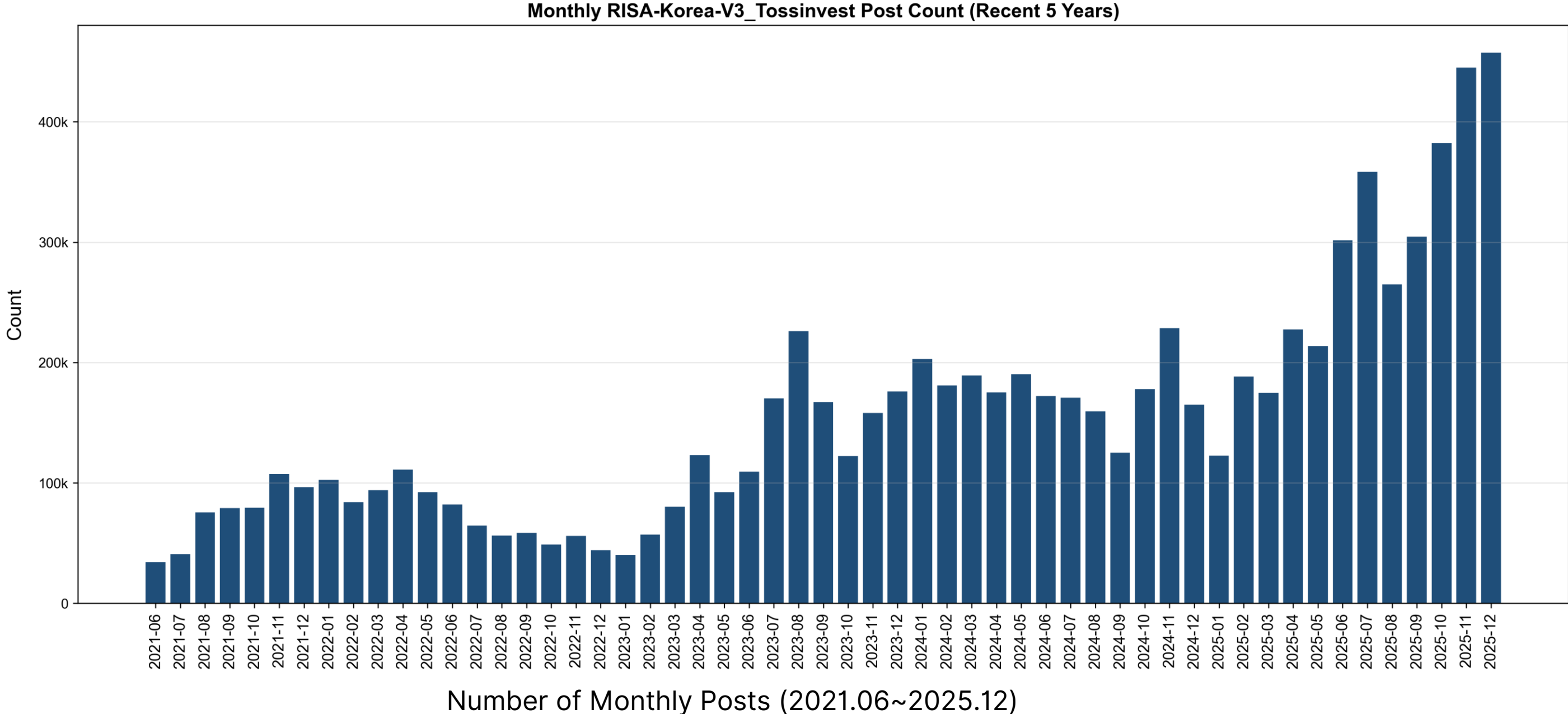


# Trend of Published Posts -Naver

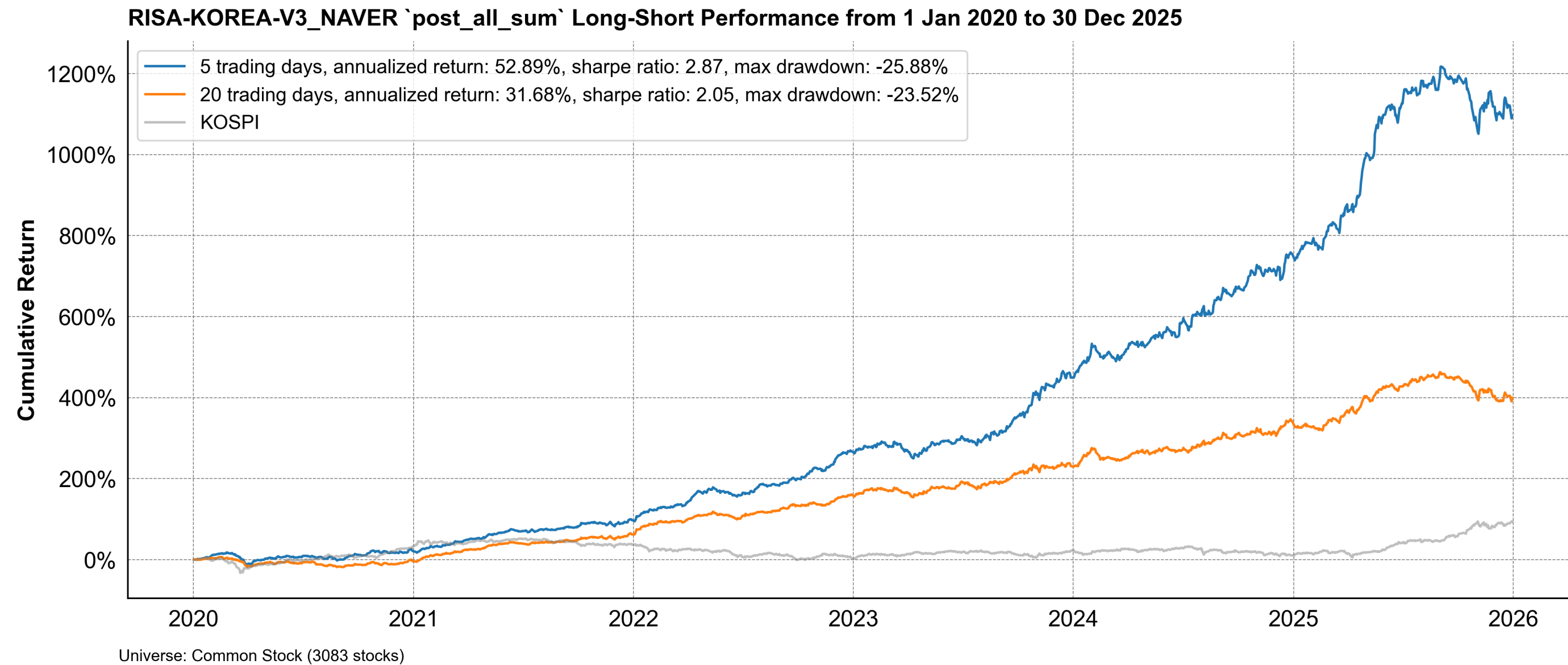


Number of Monthly Posts (2021.01~2025.12)

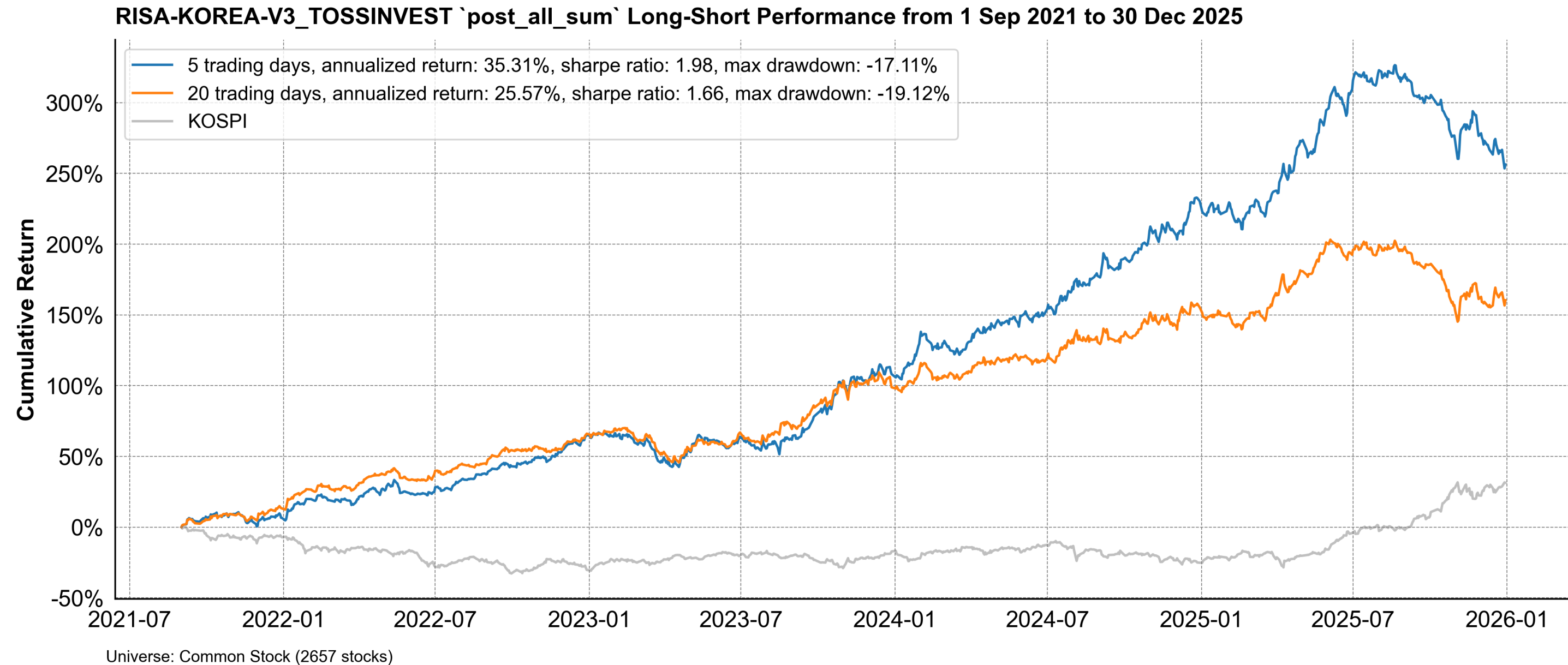
# Trend of Published Posts -TossInvest



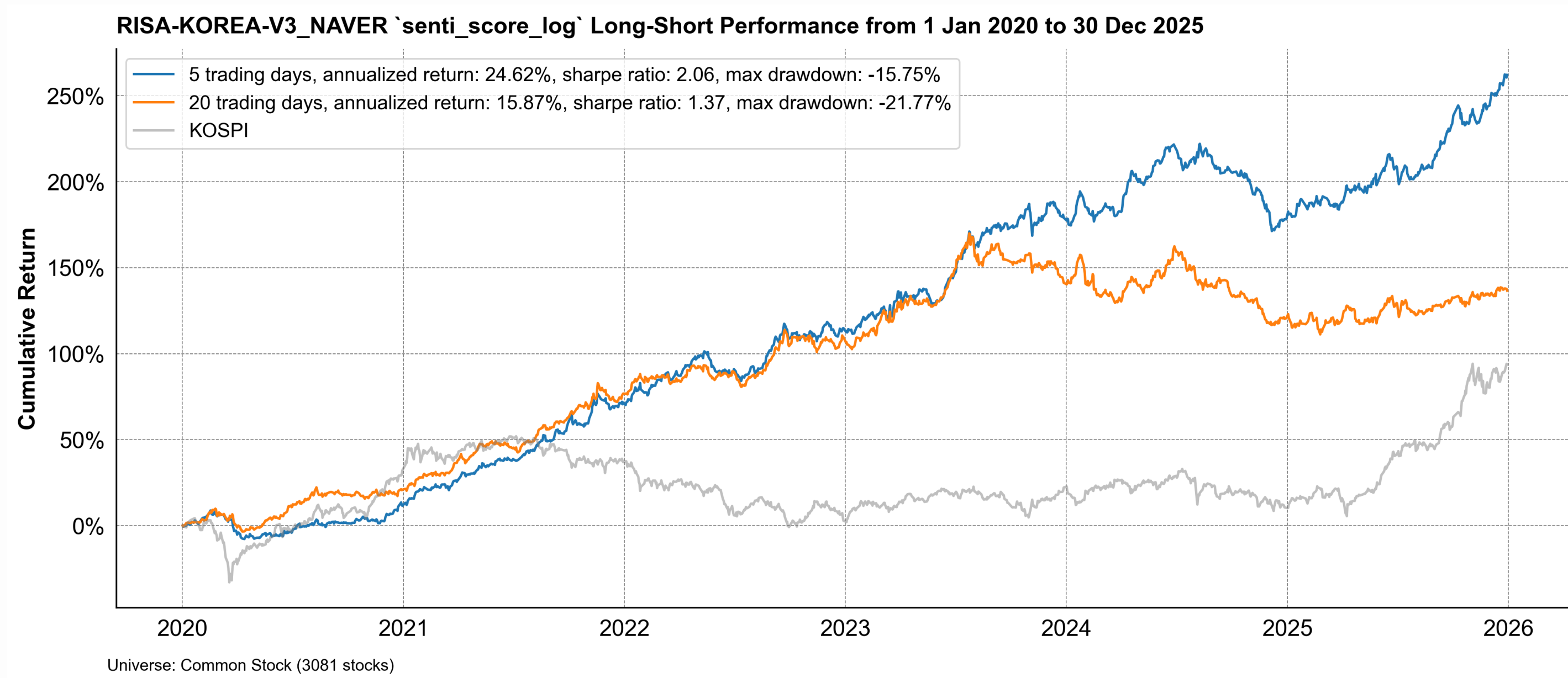
# Example of Contrarian Indicator (`post_all_sum`) -Naver



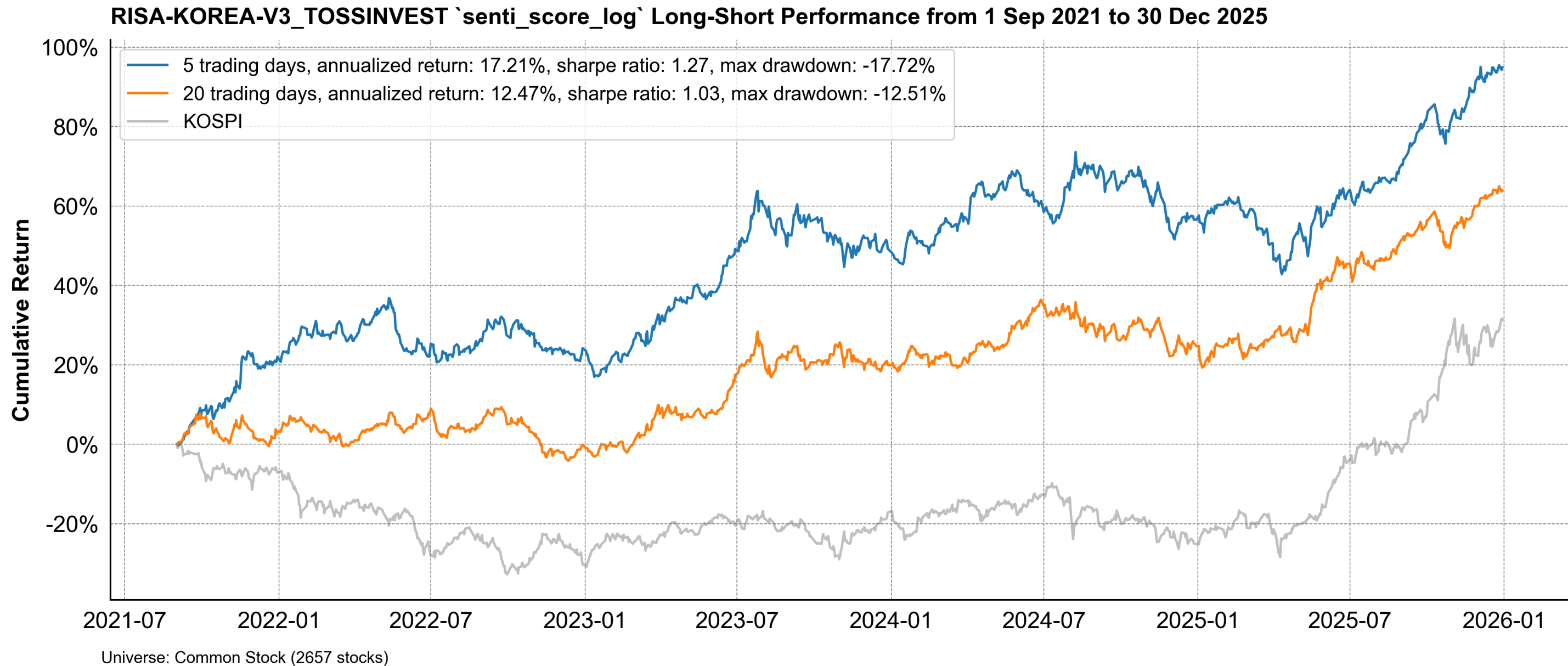
# Example of Contrarian Indicator (`post_all_sum`) -TossInvest



# Example of Positive Indicator (`senti\_score\_log`) -Naver



# Example of Positive Indicator (`senti\_score\_log`) -TossInvest



---

# Thank you!

---

Contact us:  [info@datago.com.hk](mailto:info@datago.com.hk)

 [www.datago.com.hk](http://www.datago.com.hk)

# Model Lineage of RISA-Korea

Data Field	sentiment
Base Model	BERT
Base Model Pretraining Cut-off Date	Aug 2024
Finetuned Training Data Cut-off Date	Jan 1 2018 – Dec 31 2023