



# TWSE Short Sell Quota Scraping Task

Edrick Lam  
Dec 4, 2022



# Introduction

---

- TWSE short sell quota – the available quota for placing short position via security lending
- Elements in report:
  - Explain findings
  - Analysis on significance of this real time update and its implication on daily trading work
  - Explain your methodology on scraping the data
  - How would you design the alert system to detect significant event
- Optional elements
  - Trading ideas around the info
  - Knowledge around TWSE Micro-structure
  - Your struggles in working on the project

# Findings, Analyses and Trade implications


# Looking at the data...

## Data description

- Due to time constraint, the analyses are based on the dataset of real time short quota collected from **11/29/2022 to 12/2/2022** covering **1677** stocks listing in both TWSE and Gretai.
- Some data entries are being dropped in particular metrics due to fail to capture the transactions before the program runs (8:55am). Notes are added to the related analyses.

## Example

- 0050.TW on Nov-29 will not be considered in the **calculation of overall quota consumption** velocity since there are transactions carried out before the program started scraping.



request_date	request_time	stkno	txtime	slblimit	ts
2022-11-29	08:53:12.717576	0050	08:46:18.94	430	8461894
2022-11-29	13:28:40.712057	0050	13:27:56.92	241430	13275692

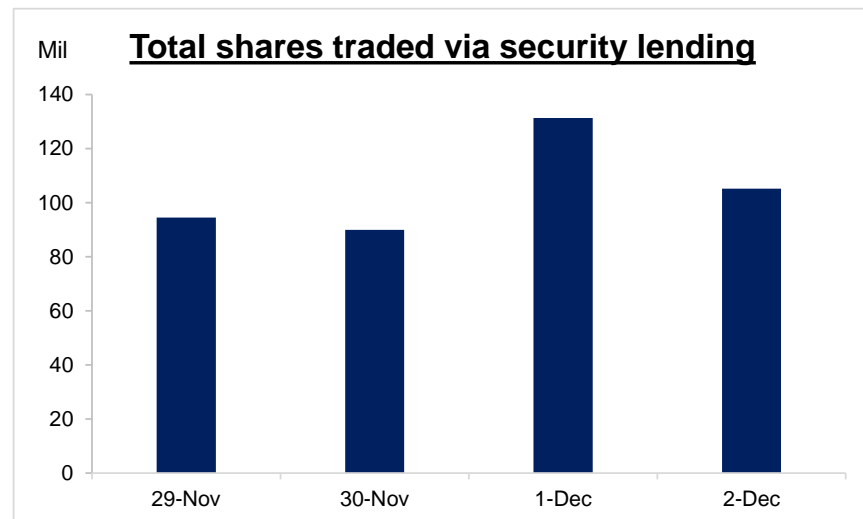
# 1. Larger turnover in month-start

- **Finding:**

- Larger turnover (both short selling and buying back) on Dec 1, 2022, **40% higher** than the previous two days.
- Possibly due to position/portfolio rebalancing or rolling over of derivatives in new month.

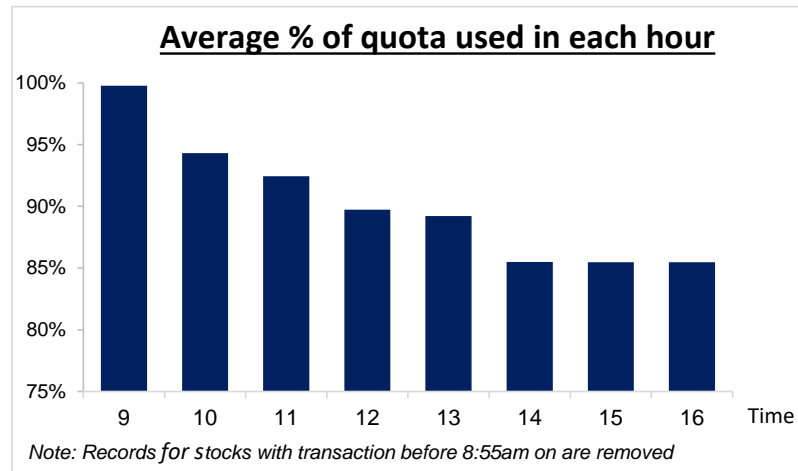
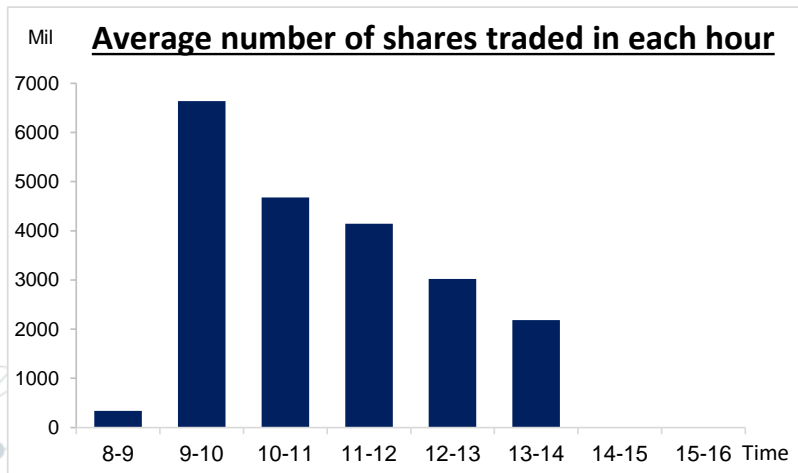
- **Implication:**

- We should expect a significantly **hotter market in start of month.**
- We should pay extra attention in subsequent alerting system since month-start would need another set of rules, e.g. a greater alerting threshold of the size of short order.



## 2. Higher trading velocity in first hour

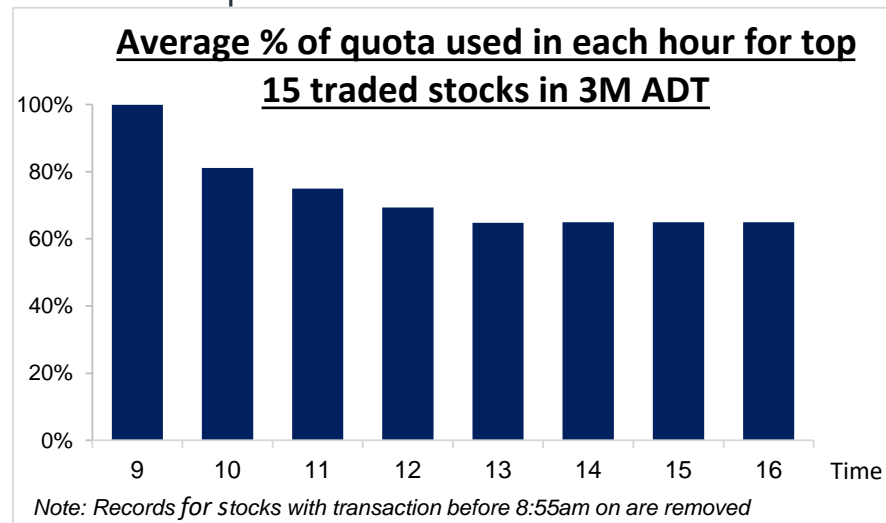
- **Finding:**
  - Most security lending turnover happens in the **first trading hour** (9-10am), which consumes > 5% of the short quota on average. Since the overall short quota drops, we can conclude that **most of the orders are short selling instead of buying back.**
  - Noticeable order in pre-trading hours period, but due to dropping of incomplete data, number is not guaranteed to be exact.
  - Since market closes at 1:30pm, we can see that the short selling is active in the last 30mins before market close comparing to 60 mins interval of other periods.
  - Muted short selling orders volume after 2pm.



# Cont'd

- **Implications:**

- We should closely monitor the market in the first trading hour, since for the top 15 traded stocks in the past 3m, **20% of the quota will be used in the first trading hour on average.**
- Which means, the team may need to consider executing pending orders or part of the order in the early hours to ensure that there will be sufficient quota.
- If the quota consumed faster than expected, maybe we should send a warning to related PM to remind them placing order if there is any, or they will need to consider other tools to short within their mandate.



### 3. Most actively traded equities via SBL

- **Findings:**
  - Quite a few stocks have their short quota exhausted almost every of the past 4 days.
  - Number of transactions are likely **positively correlated** to the short interest ratio.
- **Implication:**
  - We can **check the daily available short quota** one day in advance and estimate the use the quota. And alert the PM in advance about any execution time considerations on the equities with high quota usage.

Stocks with high quota usage in the past 4 days			
Stkno	Quota used	Stkno	Quota used
0056	100%	2609	99%
2486	100%	2006	99%
3034	100%	3036	99%
2441	100%	00900	99%
6505	100%	1303	98%
3576	100%	2539	98%
2356	99%	2881	98%
1455	99%	1304	98%
8926	99%	1718	98%
1314	99%	2412	98%

*Note: Due to data missing issue, stocks having transaction before 8:55am are removed*

Most transacted stocks in the past 4 days		
Stkno	Average transactions	Short interest ratio
2609	222	5%
00878	199	1%
006208	196	1%
00636	190	1%
4123	187	0%
00655L	182	2%
2303	179	5%
2317	176	0%
2610	168	5%
6180	162	0%
TWSE Avg	-	1%

*\*Data as of Dec 2, 2022*

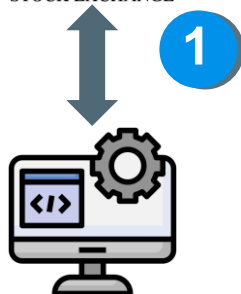
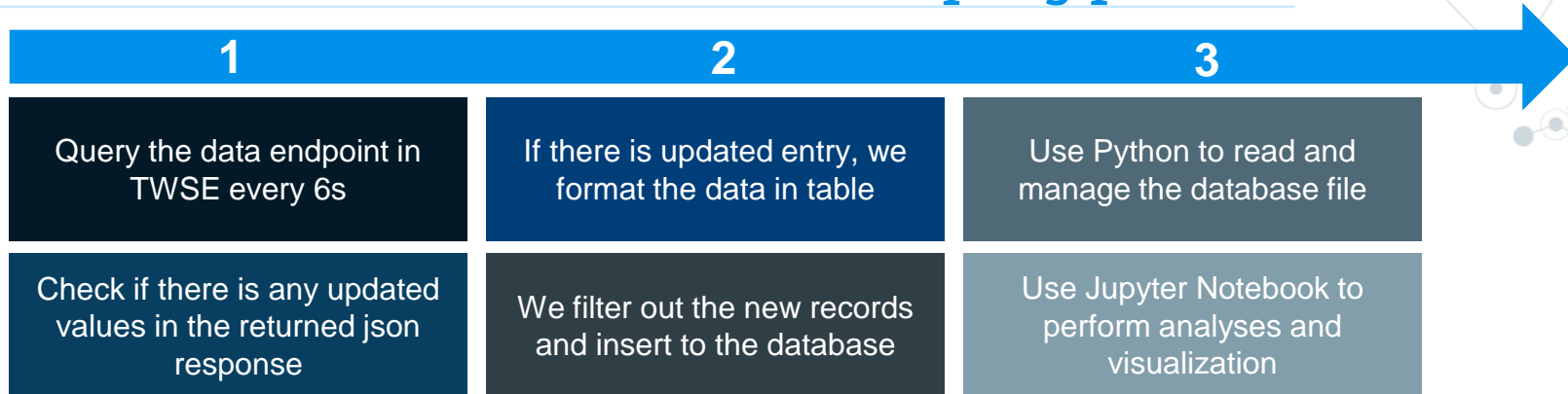




# Behind the scene: How to scrape the data?



# Overall infrastructure of the scraping process



# 1. Requesting for the data

- **Runtime:**
  - Program is run between 8:55am to 5:30pm to cover the whole trading hours.
- **Requesting endpoint:**
  - Investigating the targeted website's network traffic could find that data are updated with the following endpoint:

▼ **General**

**Request URL:** `https://mis.twse.com.tw/stock/api/getStockSbIsCap.jsp?delay=5000&_=-1670082007408`

**Request Method:** GET

- **5 seconds request interval:**
  - Observing the updating request, there is a 5 second delay in the endpoint query.
  - Which means that by default, TWSE web portal fetch **update and display on the website every 5 seconds**.
- If we make request every 6 seconds, there will be at most 5 seconds of delay between local database and TWSE database data

## 2. Storing data to local database

- **Checking Update:**
  - Local database is controlled by a .json object
  - We check for existence of new data using **MD5 hashing**, which will return a significantly different value for tiny changes in data.
  - The json control object contains the latest hash, if the latest data response has a different hash, program knows that there is new data.
- **Inserting data:**
  - New data records are detected using the following condition:
    - If there is **no record having the same request date, same last modified time and same stock number**, then the record is **considered as a new record**.
  - In other words, a **new record could have changes in last modified time solely with no change in security lending short selling limit**.
  - With this definition, we would be able to capture all the updates, including those with no changes in short selling quota, but only changes in the last modified time.  
We insert the new data records into an **SQLite database**.

# Code snippets – 1

([GitHub link](#))

## The main scraper class

```
class ShortQuotaScraper:
    """
    The main TWSE Short Quota Scraper
    """
    def __init__(self, db_path=r'../data/TWSE_SQ.db'):...

    def scrape(self, data_key='twse_sq'):...

    def check_duplicated(self):...

    def read_control_dict(self):...

    def _export_control_dict(self):...

    def load_db_engine(self):...

    def request_data(self, data_key):...

    def check_hash_update(self, data_key, new_hash):...

    def save_hash(self, data_key, new_hash):...

    def push_frame(self, df_to_push:pd.DataFrame, data_key, tb_type:str):...

    def _query_new_data(self, df_new_tb, data_key):...

    def export_frame(self):...
```

## The response handler class

```
# Implementing class
class TwseResponse:
    """
    Class to handle the twse response object
    """
    def __init__(self, res, res_time):...

    @staticmethod
    def check_response_status(res):...

    @staticmethod
    def request_data_cleaner(res):...

    def create_dataframe(self):...

    def create_meta_dataframe(self):...

    def get_hash(self):...
```

# Code snippets – 2

([GitHub link](#))

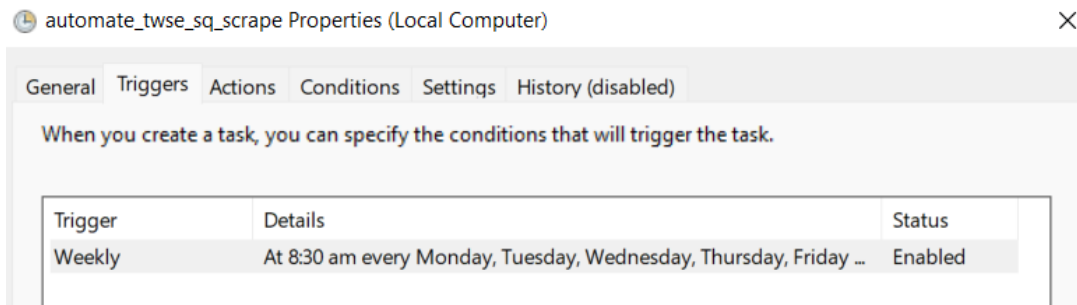
## The main script driving the scraping process

```
if __name__ == '__main__':  
    # Define the scraper  
    db_path = os.path.join(MAIN_DIR, 'data', 'TWSE_SQ.db')  
    tsqs = ShortQuotaScraper(db_path)  
  
    # Assistant variables  
    request_count = 0  
  
    # Scraping data  
    while True:  
        now_time = datetime.datetime.now().time()  
        try:  
            if now_time >= trading_hour['start'] and now_time <= trading_hour['end']:  
                print('-'*50)  
                tsqs.scrape('twse_sq')  
                request_count += 1  
                time.sleep(6)  
            else:  
                print(f"{PromptType.SYS.value} Stopping scraper, total request sent is {request_count}...")  
                break  
        except Exception as err:  
            print(f"{PromptType.ERROR.value} Encountered error: {err}, retry...")
```

# Code snippets – 3

([GitHub link](#))

## The automation infrastructure – task scheduler



## Batch script that execute the main scraping script

```
twse_sq_scrape.bat - Notepad

File Edit View

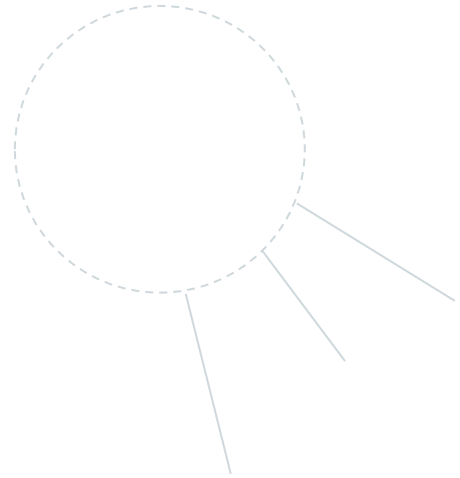
REM Author: Edrick Lam
REM Create date: 11/27/2022
REM Automated Batch Script to drive the scraper

SET project_dir=C:\... \twse_sq_scraper
CALL %project_dir%\venv\Scripts\activate.bat

CD %project_dir%

CALL %project_dir%\venv\Scripts\python.exe %project_dir%\main\dataScraping.py"
PAUSE|
```

# Designing the alert system





# Designing a short quota alert system

- To achieve real time monitoring, we can simply run the scraping script with another script with the monitoring criteria.
- Therefore, when the system detected alerting condition in the database, we can alert the team by emails or prompting messages.
- Ideas of real-time alert:
  1. Opening quota alert
  2. Quota consumption velocity alert
  3. Large movement in quota alert
  4. Overall market short ratio alert
  5. Consecutive session alert

# Designing a short quota alert system

## 1. Opening quota alert

- Setting:
  - The short sell quota of each stock is set one day in advance according to the average volume traded.
- Alert:
  - We can alert the team once the quota is below the rolling average by 1SD (~16% probability) and if the stock has a recent history of high quota consumption.
- Follow-up action:
  - Notify the PMs to be aware of position closure issue.

## 2. Quota consumption velocity alert

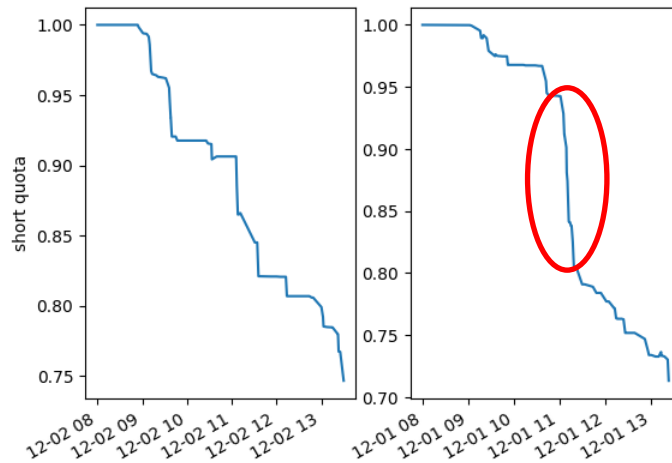
- Setting:
  - We can track the historical hourly quota consumption speed and record them down as reference.
- Alert:
  - Alert the team when a stock's consumption velocity
- Follow-up action:
  - Notify the PMs to be aware of position closure issue

# Designing a short quota alert system

## 2. Quota consumption velocity alert

- Setting:
  - We can track the historical hourly quota consumption speed and record them down as reference.
- Alert:
  - Alert the team when a stock's consumption velocity appears as faster than the historical average and below the historical average level of a particular hour. (acceleration of short selling)
- Follow-up action:
  - Check the current pending orders to see if there are trades that need to be prioritized in order to fulfill.

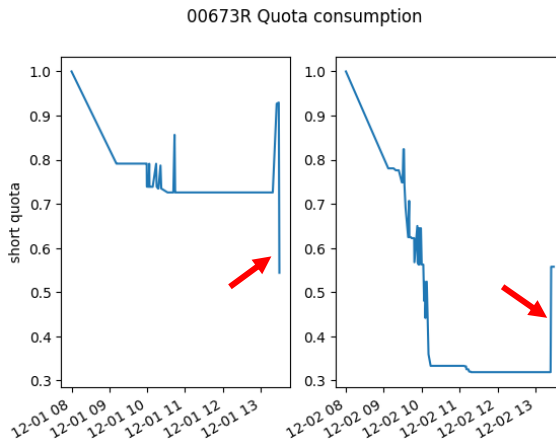
2330 Quota consumption



# Designing a short quota alert system

## 3. Large movement in quota alert

- Setting:
  - Since the data are updated when there is transaction, we can monitor the large movement and give this info to the team.
- Alert:
  - Alert the team if the share transacted in a single update is beyond the top 15% situation in the 30D history.
- Follow-up action:
  - Check for news update and research to see if there is any reason for the moves.
  - Estimate if this will cause any execution difficulty on the next day



# Designing a short quota alert system

## 4. Overall market short ratio alert

- Setting:
  - Overall quota usage reflects the market sentiment (especially foreign institutes).
- Alert:
  - Alert the team when the average quota consumption % is higher than the 30D history at the same time.
- Follow-up action:
  - Check for news and let the team and PMs knows after the end of day.

## 5. Consecutive session alert

- Alert:
  - Alert the team when a stock has consecutively deviate from recent history in quota consumption.
- Follow-up action:
  - Let the team know what we can expect on the next trading session to handle the execution with this information.

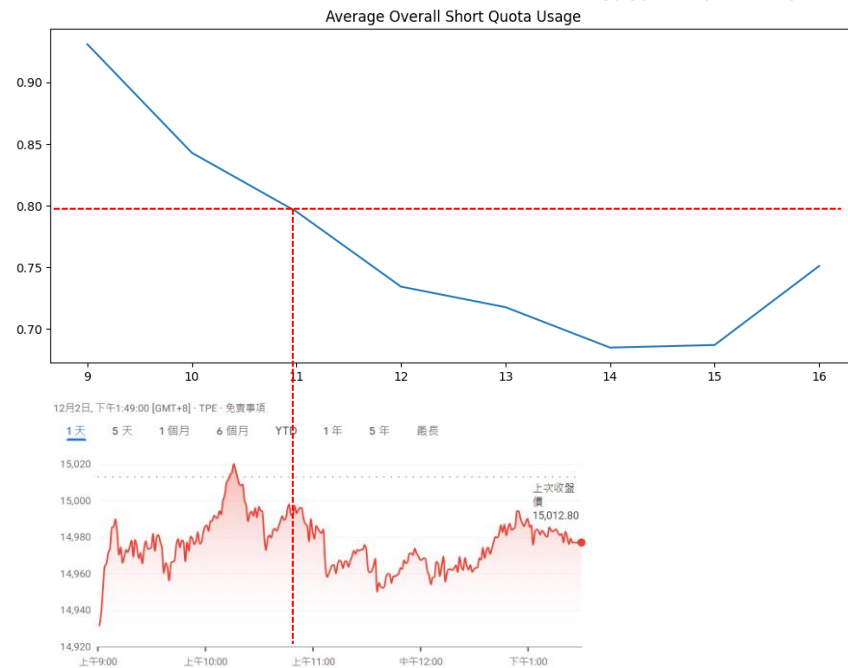
# Trading ideas

# Trading ideas: Short Index

## 1. Short index

- When the intra-day market overall short ratio lower than 80%
- The quota usage drops in the previous hour as well
  - => short the TAIEX.
- When the hourly quota usage did not drop
  - => close the positions

Notes: More frequent data and dates will be needed to decide the exact threshold



Source: Google Finance

# Trading ideas: Long short first-hour quota

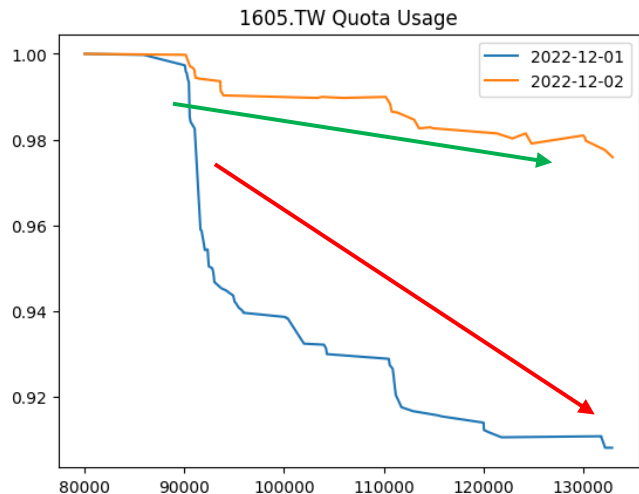
## 2. Long short

- Long the stocks with higher than recent average first-hour quota consumption for the day.
- Short the stocks with lower than recent average first-hour quota consumption for the day.
- Rationale:
  - Short sell quota and price tends to move in a positively correlated direction.
  - The first-hour short sell quota consumption reflects the market sentiment and somehow has a guidance effect on the subsequent intra-day movement.
- Note:
  - Due to incomplete data, ideas are generated only upon the complete records for the past 4 trading days.



# Short quota consumption reflects on price

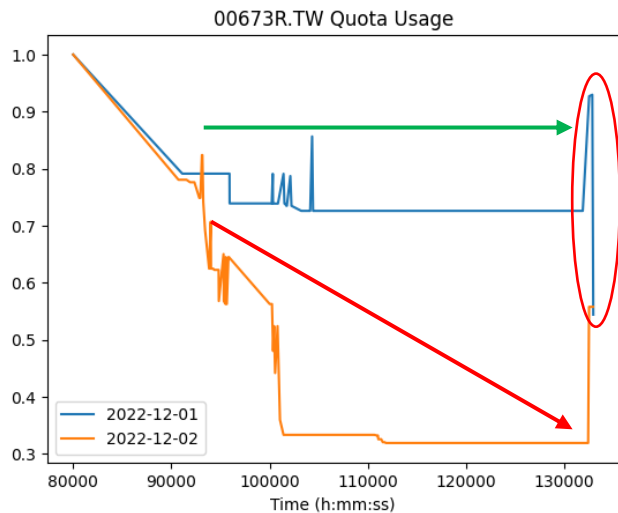
- Date with fewer usage in first trading hour perform better than date with relatively high usage of quota in the first trading hour.



Source: Yahoo Finance

# Cont'd

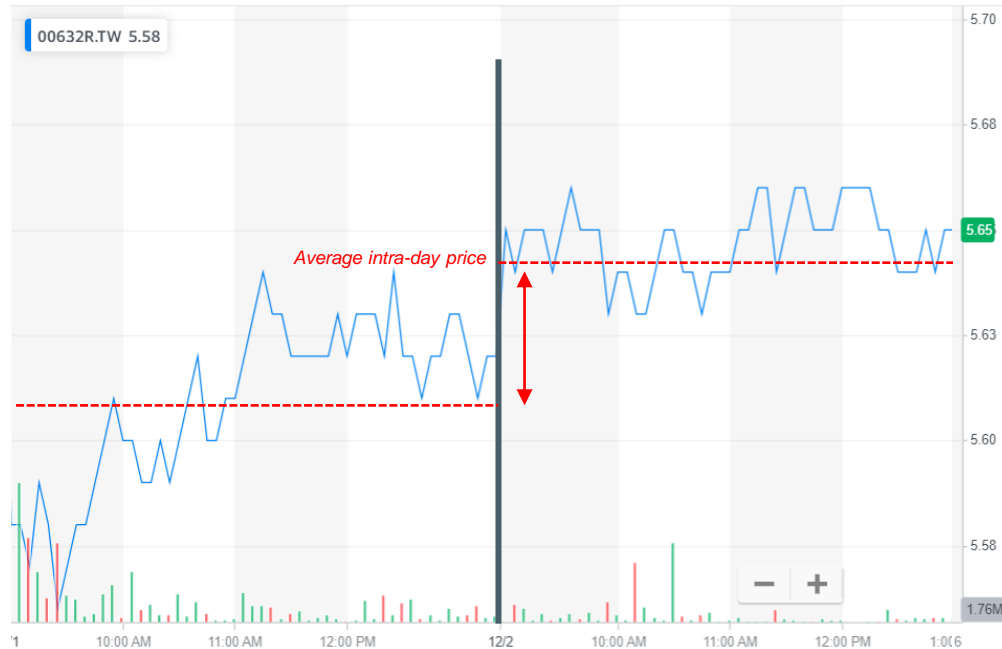
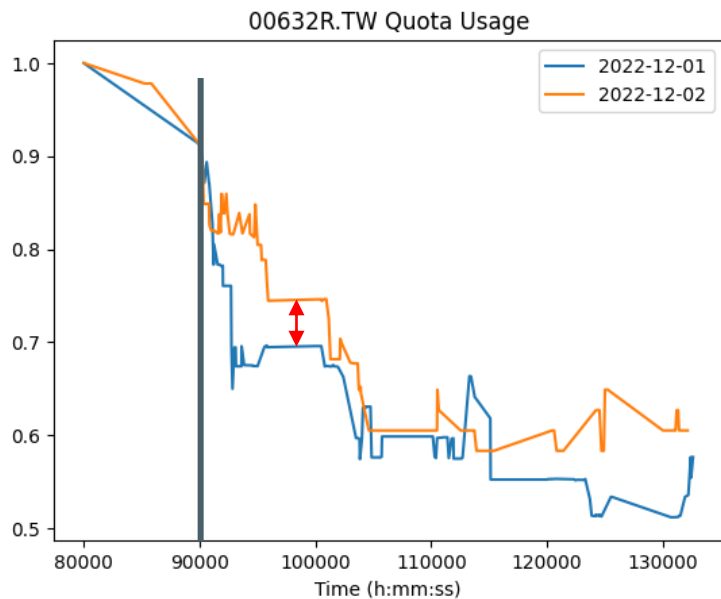
- The diverged short quota consumption movements reveals the diverged price movements.
- The short quota usage in first trading hours having a guidance on the subsequent trend.



- **Notes:** The example equities is *S&P GSCI Crude Oil 1x Inverse ER Futures ETF*, so the movement could be impacted by the underlying asset, yet it reflects somewhat the correlation of short quota consumption and price movement.

# Cont'd

- The cross-day prices and short quota usage are moving steadily together with the consumption.
- Dec 2 has a better performance in line with the overall less consumption in short sell quota.

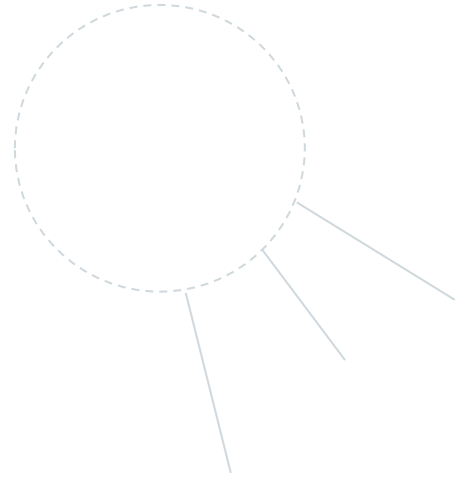


# Other ideas

---

3. High short interest ratio = potential short squeeze
  - Stocks with high short interest ratio may encounter short squeeze
  - Price surge with short quota consumption quota increase may be a signal of short squeeze and short positions are being closed. Resulting another round of price surge.
4. Short selling % as of ADVT
  - Stocks with less liquidity may be more suffered from SBL short selling

# TWSE Micro-structure and Struggles



# Knowledge around TWSE Micro-structure

## 1. Recent update on TWSE Short Quota

- Intraday security lending short selling quota are calculated as a percentage of average volume over the previous 30 trading sessions.
- There are two adjustments announced in Oct 2022.
  - Oct 1: 30% of 30D ADVT -> 20% of that
  - Oct 11: 20% of 30D ADVT -> 10% of that

## 2. SBL Quota reflects institute investors' sentiment, mainly foreign funds

- Local retail investors would have higher flexibility to short sell with broker
- Due to the forex rate protection, foreign fund are restricted to invest within one week, so foreign institutes preferred to participate the market via security lending.

# Struggles with the projects

---

- Data issue
  - Ambiguous time to start scraping:
    - Earliest transaction happens far before the market open
    - We should scrape from 5 mins before order placing time (8:30am)
  - Small sample size
    - 4 days of data, some stocks are with incomplete records makes it hard to do robust analyses
- Not enough understanding towards the market landscape
  - Understanding confusing terminologies given time constraint
    - SBL Short Sales Balance
    - Available volumes for SBL short sales
- Lack of trading experience
  - Do not have a sense on the needs of actual trading execution

# Useful links

---

- [Real Time Available Volume for SBL](#)
  - [Available volumes for SBL short sales](#)
  - [Balance of securities provided as Financing Collateral](#)
  - [Daily Short Sale Balances](#)
  - [Historical transaction inquiry](#)
- 
- Reference:
    - [Taiwan further tightens short selling rules amid equity volatility](#)





# End

Thank You!

