**AfricaAI: Runbook - SHFEMR – SHFE**

**Summary** We want to capture margin changes from the daily Shanghai Futures

Exchange report.

**Source name** Shanghai Futures Exchange - SHFE

**Link to the source** <https://www.shfe.com.cn/publicnotice/notice/>

**Dataset name**  Shanghai Futures Exchange Margin Requirements - SHFEMR

**Frequency** Weekdaily

**SLA** Check data availability every day before 7AM EST time.

Expected Time **NA**

Instructions:

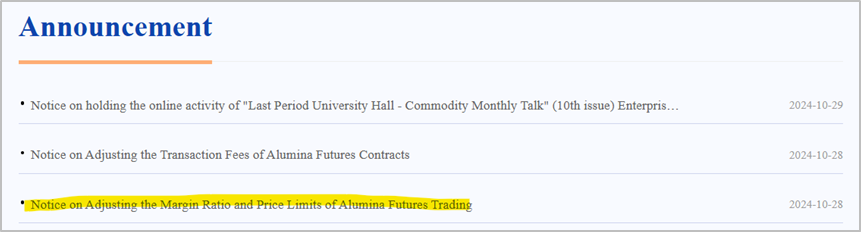
1. NB! This data should be processed only via your AWS workspace.

Go to the link: <https://www.shfe.com.cn/publicnotice/notice/>

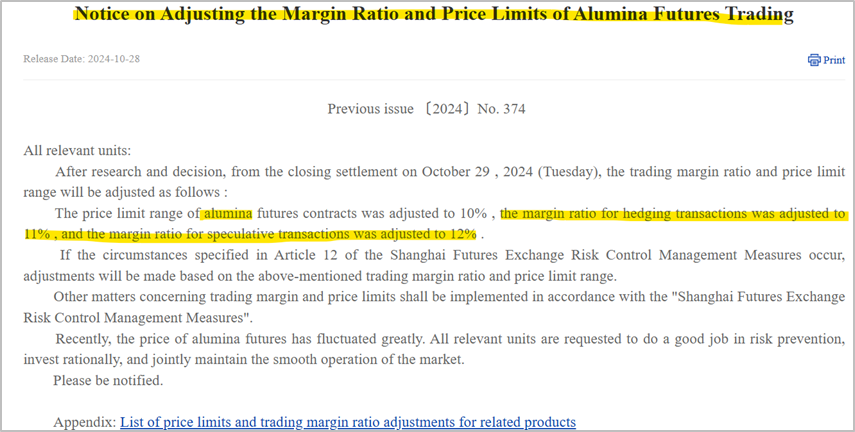
Translate the site to English from Chinese.

2. Find the most recent report with a title that begins with “Notice on Adjusting the Margin Ratio and Price Limits of …”. For example: “Notice on Adjusting the Margin Ratio and Price Limits of Lead and Other Futures Trading”.

NOTE: Sometimes required data may be published in reports with a different name, so check all reports.



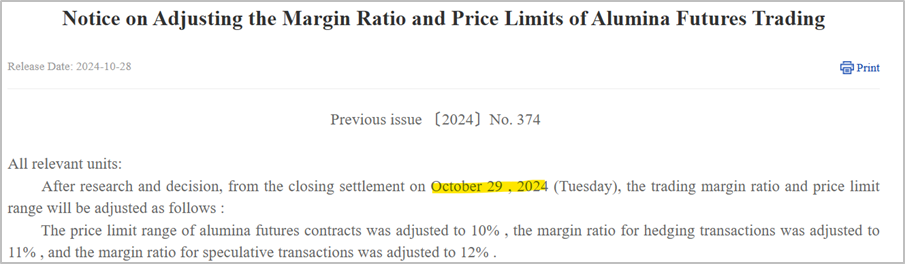
3. Open the report and find the part in the text that describes margin changes for commodities:



4. Save the appropriate margin ratio values for each of commodities and transaction types.

If the transaction type is not specified, save the value for both types: for hedging transactions and for speculative transactions.

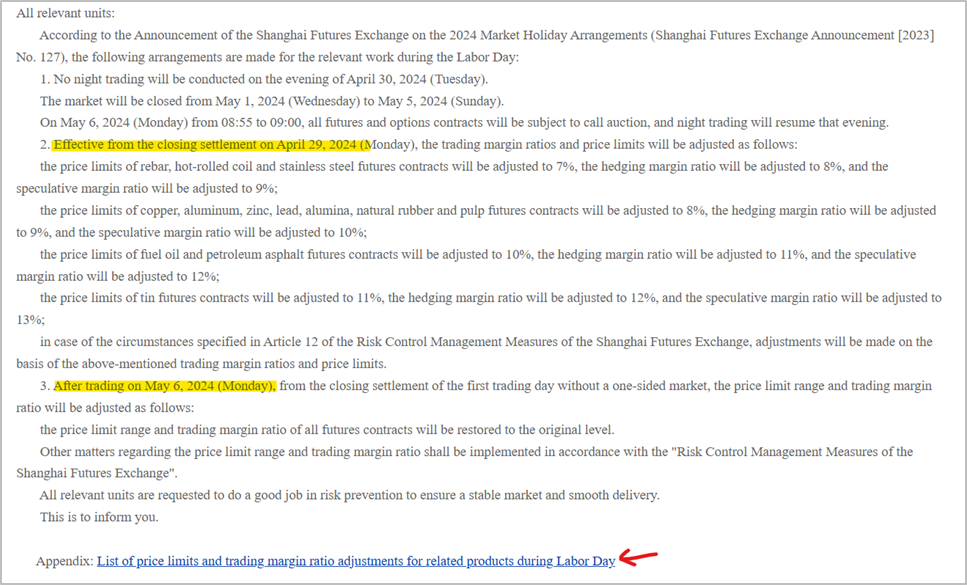
Use time period from the text - the date the release says the change is effective after the close of:



And convert this date to YYYY-MM-DD format (for example above it will be – 2024-10-29).

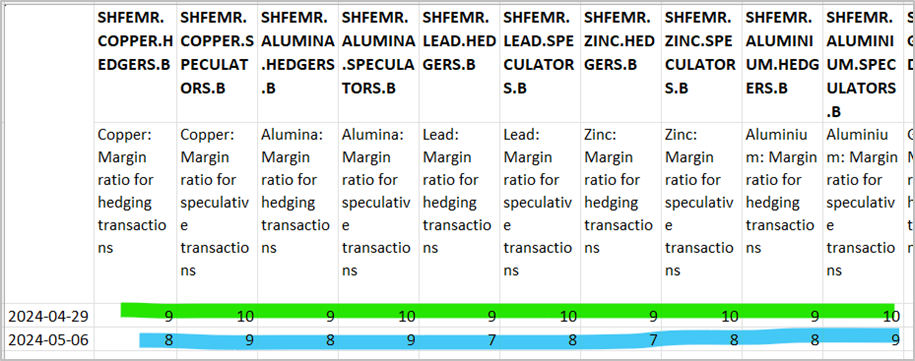
Note that there may be more than one date in the text. In this case, save the values for both dates (first for the earlier one, then for the later one). To understand which values are used in which period, you can use the file in the appendix (download \*docx file and translate it).

Below are examples of the report for 2024-04-25 and tables of values from the appendix.

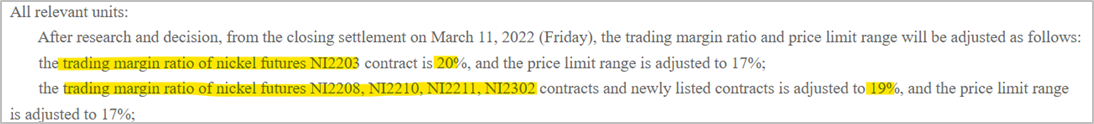




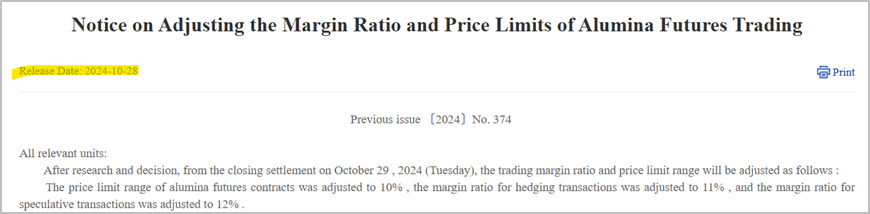
The values for this report in the DATA file will look like this (using copper, aluminum, zinc, lead, alumina as an example):



If contracts and several values for margin ratio on one date are specified for one commodity, take the average value. For example, the value will be 19.5:



**Use LAST RELEASE DATE from the report and covert to YYYY-MM-DDT11:00:00 format**.



**NOTE: Remove the "LAST RELEASE DATE" column from the META file after the initial ingest.**

Please inform the DataOps and DataEngineering teams in case of changes in the source.

Requirements to the data provided by AfricaAI

Data files format

For on-boarding new datasets and processing new releases the following requirements are applied:

• there should be one archive in .ZIP format that contains 2 files provided in .XLS format per one release - one file with DATA and one file with METADATA (both files are mandatory). Archive name should be as "DATASET\_YYYYMMDD.ZIP"", where "DATASET" - specified abbreviation for a dataset, "YYYYMMDD" - timestamp when the file was created, where YYYY - year, MM - month number, DD – day;

• the files format is strictly defined and cannot be changed by AfricaAI team.

File with DATA

Template for file name: “DATASET\_DATA\_YYYYMMDD.xls”, i.e.” MEXPENFND\_DATA\_20220601.xls”

Requirements to the file with DATA:

1. format - XLS, encoding - Unicode (UTF-8);

2. file name: should follow pattern "DATASET\_DATA\_YYYYMMDD.xls", where "DATASET" - specified abbreviation for a dataset, "YYYYMMDD" - timestamp when the file was created, where YYYY - year, MM - month number, DD - day;

3. file content:

a. for data with daily/ weekdaily frequency: YYYY-MM-DD, where YYYY - year, MM - month, DD – day;

b. for data with weekly frequency: YYYY-WW, where YYYY - year, WW - week number (should be two-digits, i.e. 2022-01, 2022-02, etc.). Week number is defined based on the reporting date: reporting date is considered the end of the week, i.e. if reporting date is Wednesday the week is calculated from Thursday till Wednesday – based on that rule the number of the week is defined;

c. for data with monthly frequency: YYYY-MM, MM - month number (should be two-digits, i.e. 2022-01, 2022-02, etc.);

d. for data with quarterly frequency: YYYY-QX, where YYYY - year, X - quarter number;

e. for data with annual frequency: YYYY, where YYYY – year;

f. if provider specifies for reported period that data is Not Available (that could provided in corresponding notes or other way) - fill in that data point with "NA" value;

g. if provider does not specify that data is Not Available and report Empty value for reported period - leave Empty cell;

h. if provider skips some of reporting period - skip it also in the file with data as well;

i. data is provided on one sheet only (multiple sheets are forbidden);

j. two-dimensional format where rows represent reported periods and columns contain timeseries values;

k. reported period(s) should be provided in the following format:

l. values to be provided in decimal format with "." as delimiter (number of digits after "." is limited only by provided value by the source). In order to interpret data correctly the following requirements are applied:

m. column names to be provided in two first rows: in the first row - timeseries ID (CODE), in the second row – timeseries description (description in DATA file is identical to description in file with METADATA).

NB! Timeseries with different frequencies could be provided within the same file as there are different patterns for providing reported periods

File with METADATA

Template for file name: “DATASET\_META\_YYYYMMDD.xls”, i.e.” MEXPENFND\_META\_20220601.xls”

Requirements to the file with METADATA:

1. format - XLS, encoding - Unicode (UTF-8);

2. file name: should follow pattern "DATASET\_META\_YYYYMMDD.xls", where "DATASET" - specified abbreviation for a dataset, "YYYYMMDD" - timestamp when the file was created, where YYYY - year, MM - month number, DD - day;

3. file content: specified metadata attributes that are static unless provider changes something or new timeseries should be added. In case of new timeseries should be added/ updated/ changed - escalate that to Data Engineering team Special BWAT-DFAC Data Engineering <SpecialBWAT-DFACDataEngineering@epam.com> at and Cc: Data Operations team at Special BWAT-DFAC DataOps <SpecialBWAT-DFACDataOps@epam.com>

Notes on METADATA fields from file “DATASET\_META\_YYYYMMDD.xls”

1. With every data release attach file with METADATA

2. With every release change only on attribute in the file with METADATA in column “NEXT\_RELEASE\_DATE” – set “YYYY-MM\_DDThh:mm:ss” (YYYY – year, MM – month, DD – day, hh – hour, mm – minutes, ss – seconds. Time should be provide in UTC time) to the time when the next release is going to be processed.

Saving ZIP archive on SFTP

Save the releases in specified per “DATASET” folder (you may create new folder for new DATASET in case of absence of such folder).

Details on SFTP are the following:

1) Dev env

Host: africaai.datadev.dfac.io

2) PRODVAL env

Host: africaai.prodval.data-factory.io

3) PRODHA env

Host: africaai.prodha.data-factory.io

Port: 22