Session 6 –

How to QA the Metric Results for Existing Markets on MQD

CMCRC
ASSURING MARKET QUALITY

Marc Bohmann
mbohmann@cmcrc.com
Capital Markets Cooperative Research Centre

28 March 2018

Table of Contents

Introduction

- We are constantly working on expanding the amount of markets we cover in MQD
- Metrics are the core output of our computations
- Errors and other problems can occur, e.g.
  - Thomson Reuters Data errors / absence of data
  - Wrong classification/qualifiers
  - Converting issues / code bugs
- To make sure we eradicate these issues, we scan each market individually visually and in WorkFlow

Recap Daily QA Routine

- The first and most crucial part of QA is the daily routine of checking the assigned market visually
- Given the amount of metrics, it is useful to check the main metrics (spreads, daily stats, etc.) daily and a variety of ‘non-core’ metrics sporadically
- Vary the time-frame, trend line, and winsorization options!

For all Quality Assurance visit:
http://mqd-qa.aws.cmcrc.com/

Recap Daily QA Routine

- The second part is checking the WorkFlow for failed and long-running jobs
- The WorkFlow calendar offers a variety of ways to check for failed jobs
- Each team member has to decide if a failed job can be handled by the business team or developers

The WorkFlow is available at:
http://WorkFlow.aws.cmcrc.com/

Existing Metric QA

- After a market is built, we need to QA the metric results in depth and frequently, this includes:
  - Downloading raw data
  - Comparing raw data to publicly available data (e.g. Yahoo! Finance)
  - Calculating metric results yourself (Excel, Python,…)
  - Confirming accuracy or reporting an issue

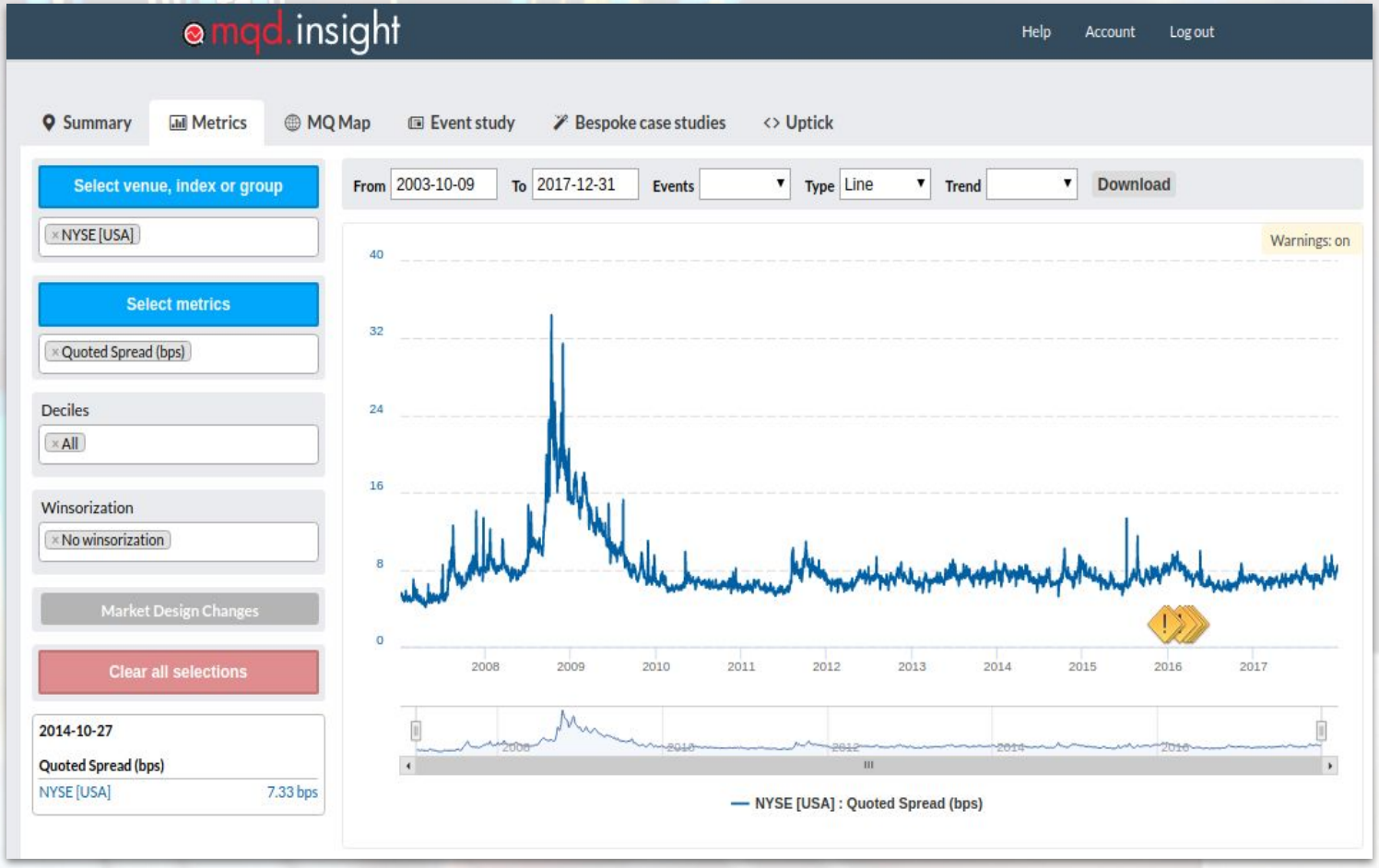
Existing Metric QA

- What follows:
  - Visually scanning a market and checking WorkFlow (metric generate and sync)
  - Opening a JIRA ticket
  - Checking one NYSE stock and downloading MQD output
  - Checking AWS S3 for results
  - Downloading raw data from TRTH
  - Comparing Close and Volume to Yahoo! Finance
  - Calculating the Quoted Spread in Excel and comparing

MQD Visual Checks

MQD Visual Checks

On first sight we can determine if the market and metric results ‘look good’. This is not a substitute for a proper QA but gives us a general impression about the markets’ health.

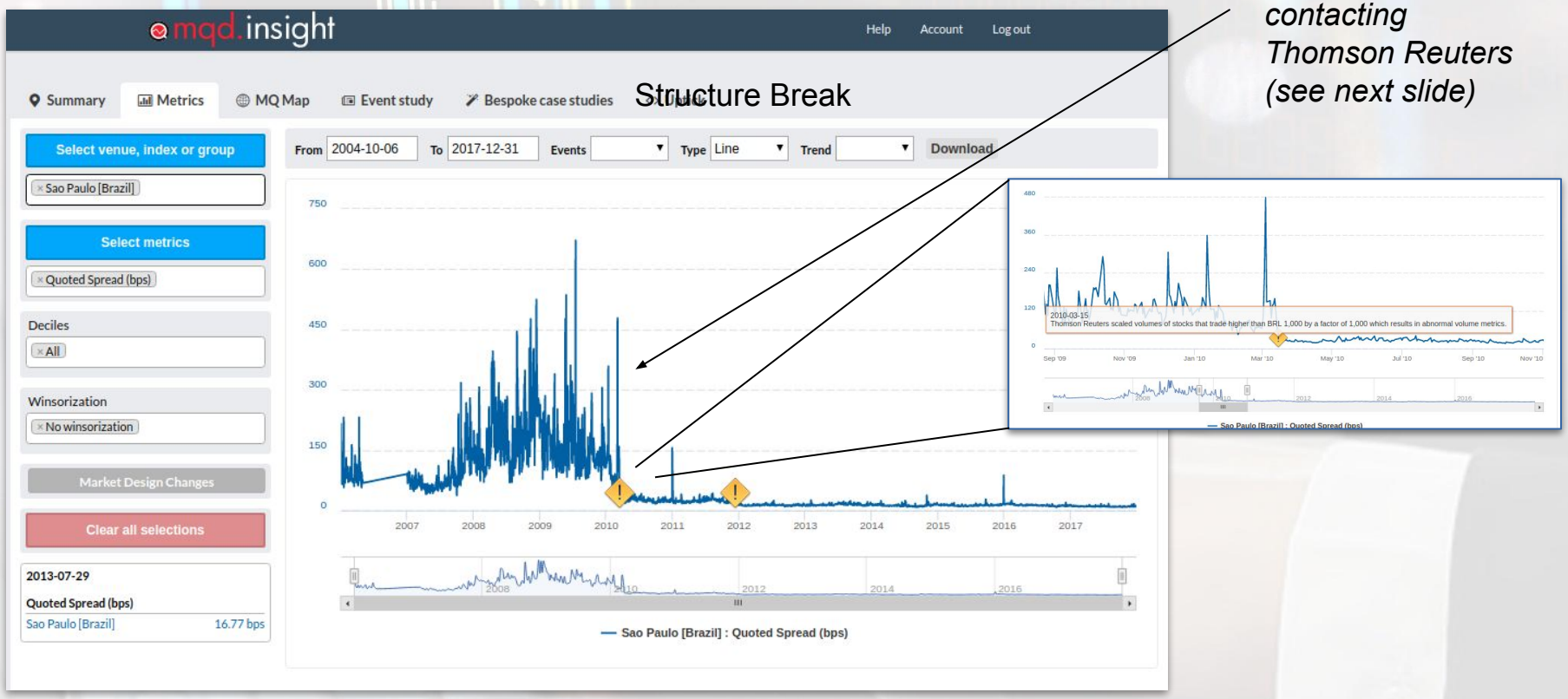


MQD Visual Checks

Market-wide example
We need to check the data for the time prior and after the structure break in TRTH

Eventually problematic metric results -> 1) Structure break

Resolving the issue included: TRTH raw data check and contacting Thomson Reuters (see next slide)



MQD Visual Checks

Market-wide example
We need to check the data for the time prior and after the structure break in TRTH

Also, the trading hours were out of sync and we needed to amend it in marketconfig.

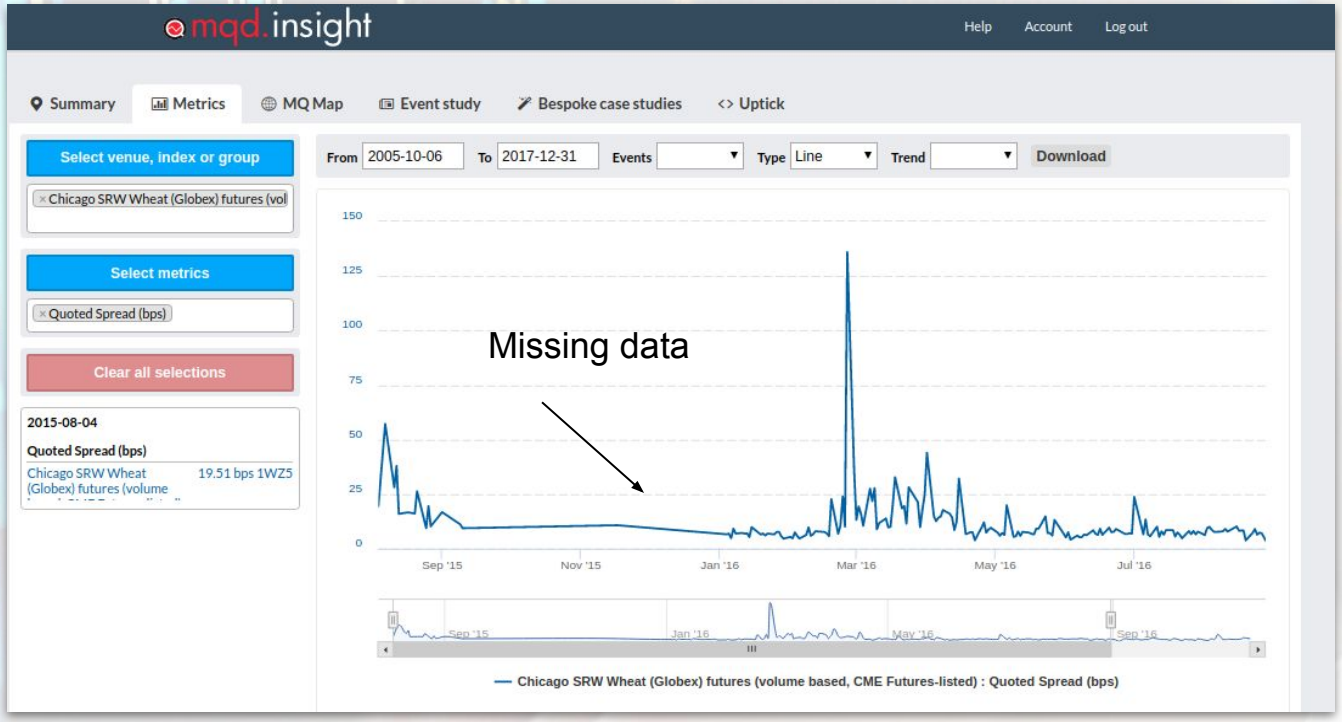
CBT_FUT.yaml | IEU_FUT.yaml | CMX_FUT.yaml | CBF_FUT.yaml | JNB.yaml | listing_markets.yaml | SAO.yaml

```yaml
sessions: #Trading hours taken from TRTH data observation
  2007-01-01:
    '11:00:00': PRE_OPEN
    '11:02:00': OPEN
    '18:00:00': CLOSING
    '18:15:00': CLOSED

  2007-03-10:
    '10:00:00': PRE_OPEN
    '10:02:00': OPEN
    '17:00:00': CLOSING
    '17:15:00': CLOSED

  2007-10-15:
    '11:00:00': PRE_OPEN
    '11:02:00': OPEN
    '18:00:00': CLOSING
    '18:15:00': CLOSED

  2008-03-18:
    '10:00:00': PRE_OPEN
    '10:02:00': OPEN
    '17:00:00': CLOSING
    '17:15:00': CLOSED

  2008-10-20:
    '11:00:00': PRE_OPEN
    '11:02:00': OPEN
    '18:00:00': CLOSING
    '18:15:00': CLOSED
```

MQD Visual Checks

Here, we had to re-convert data but did not perform a matrefresh

Eventually problematic metric results -> 2) Missing data



- Steps to locate the root of missing data include WorkFlow job include checking (matrefresh, data output), S3 raw data check, and S3 metric output check

MQD Visual Checks

Eventually problematic metric results -> 3) Spike

- Steps to locate the root of a spike are similar to missing data but include the search for an event that could have caused the spike

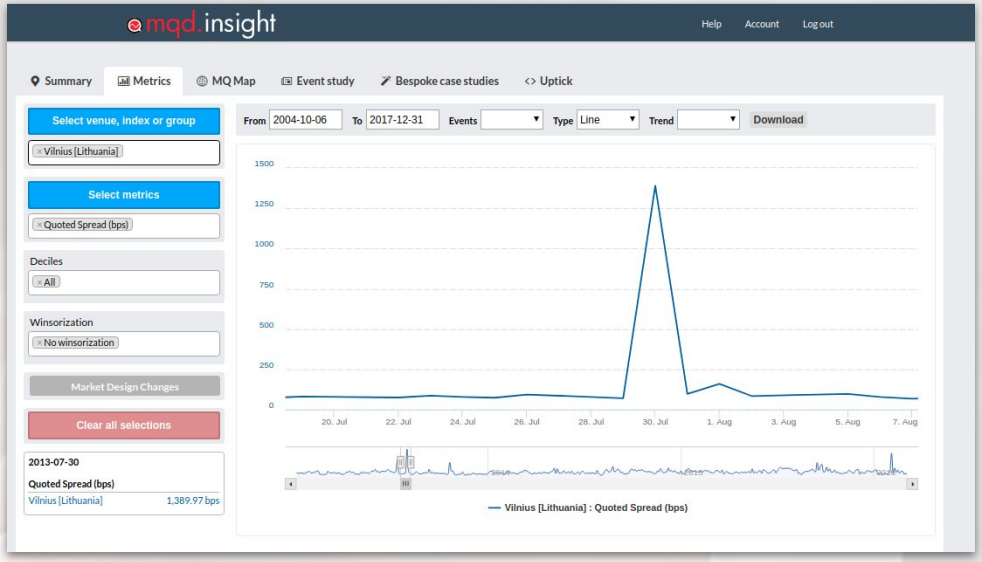
MQD Visual Checks

Eventually problematic metric results (Spike)

On Vilnius there is a spike on 2013-07-30 in quoted spread. The Quoted Spread is at 1,390bps while the average values are under 100bps.

- we investigate this spike the following way:

1. Market wide or security only?
 (Click on spike to get a top security list)
2. Events on that day
3. WorkFlow job
4. S3 metric generate output
5. TRTH raw data

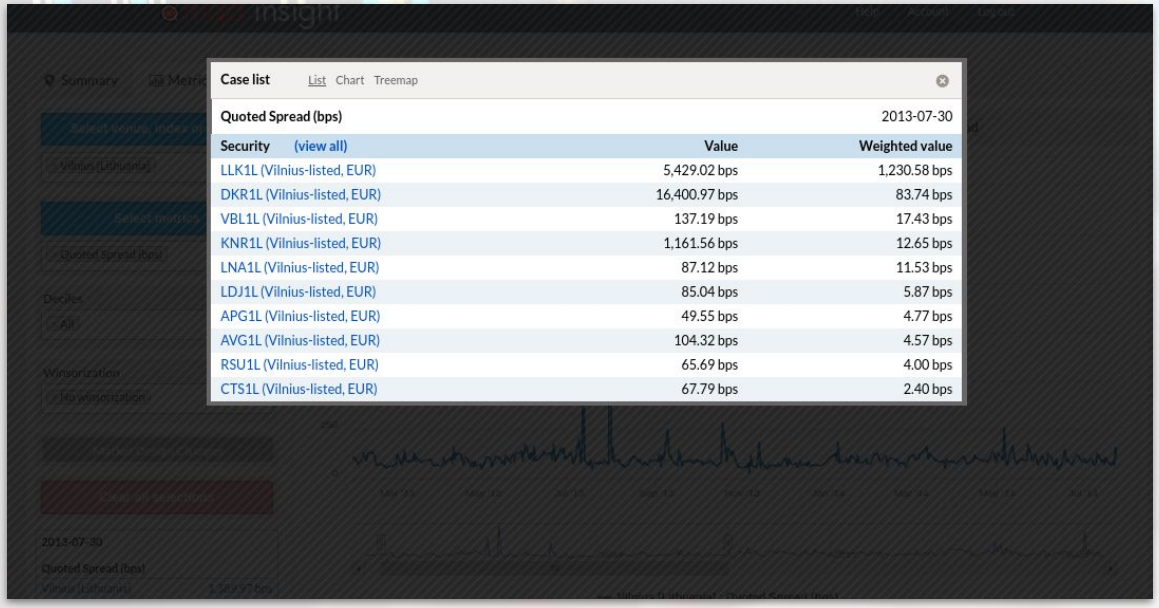


And always:
OPEN A JIRA TICKET

Let's try a live example…

MQD Visual Checks

Single security example
By clicking on the spike, we see the security list:



Case list List Chart Treemap

Quoted Spread (bps)		2013-07-30
Security (view all)	Value	Weighted value
LLK1L (Vilnius-listed, EUR)	5,429.02 bps	1,230.58 bps
DKR1L (Vilnius-listed, EUR)	16,400.97 bps	83.74 bps
VBL1L (Vilnius-listed, EUR)	137.19 bps	17.43 bps
KNR1L (Vilnius-listed, EUR)	1,161.56 bps	12.65 bps
LNA1L (Vilnius-listed, EUR)	87.12 bps	11.53 bps
LDJ1L (Vilnius-listed, EUR)	85.04 bps	5.87 bps
APG1L (Vilnius-listed, EUR)	49.55 bps	4.77 bps
AVG1L (Vilnius-listed, EUR)	104.32 bps	4.57 bps
RSU1L (Vilnius-listed, EUR)	65.69 bps	4.00 bps
CTS1L (Vilnius-listed, EUR)	67.79 bps	2.40 bps

We see that the Quoted Spread spike is caused by one security: LLK1L
A quick check tells us it is ‘Limarko’

Let’s check TRTH how trading in Limarko looks like...

Given the trading stop and the wight allocated by the enormous volume in this small market, we can exclude the day (see previous Workshops)

MQD Visual Checks

Further examples can be found in our JIRA database, for example:

(1) BUS-1245 -> Spike was caused by news in one company and hidden (other possibility: *Winsorization* by user)

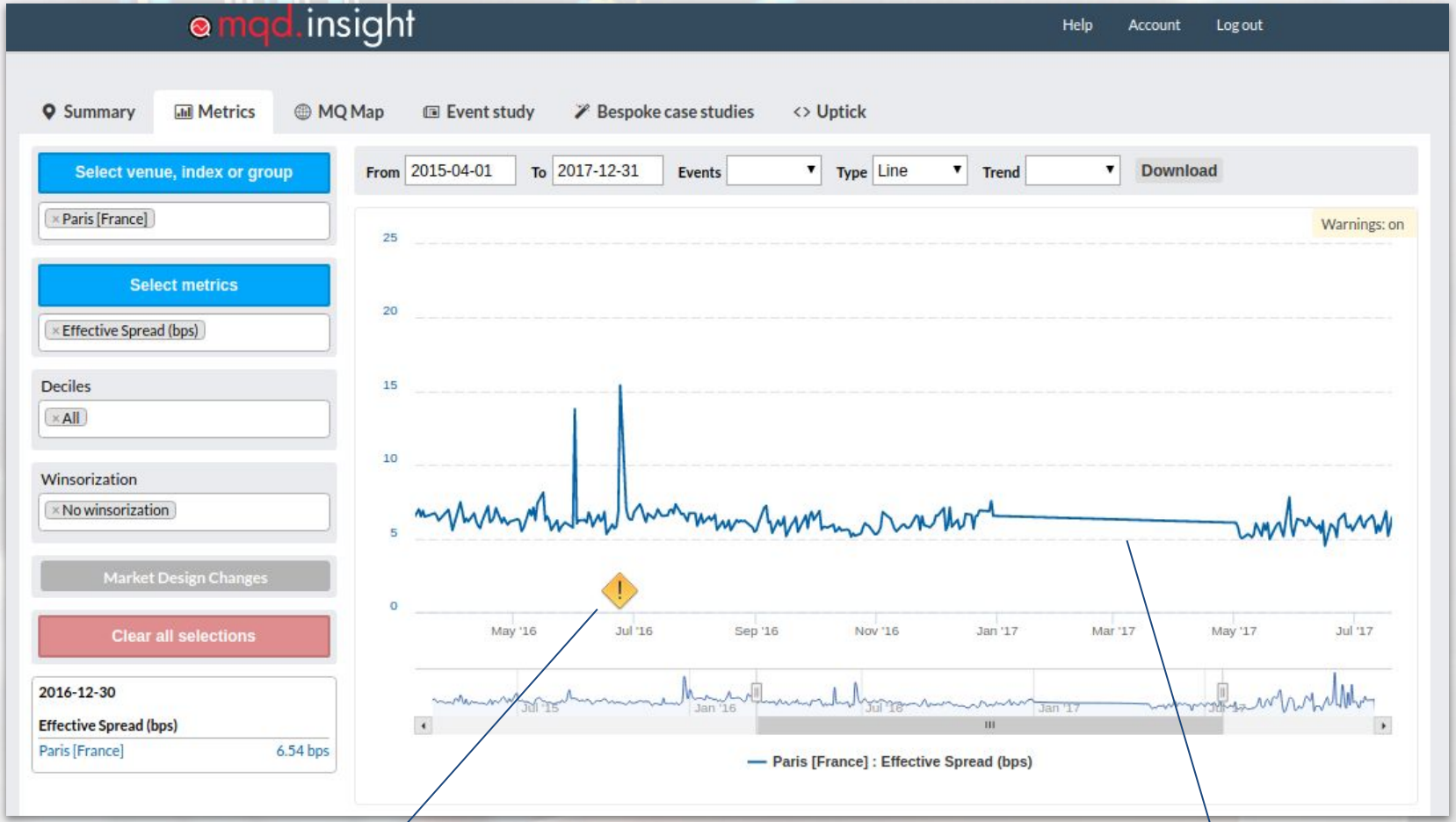
(2) DATA-1342 -> Spike was caused by a missing qualifier and fixed in the .yaml file

(3) BUS-2553 -> Structure break was caused by pretended WorkFlow jobs

(4) BUS-1692 -> Structure break due to trading hour change and fixed by amending market config file appropriately

MQD Visual Checks

Further examples

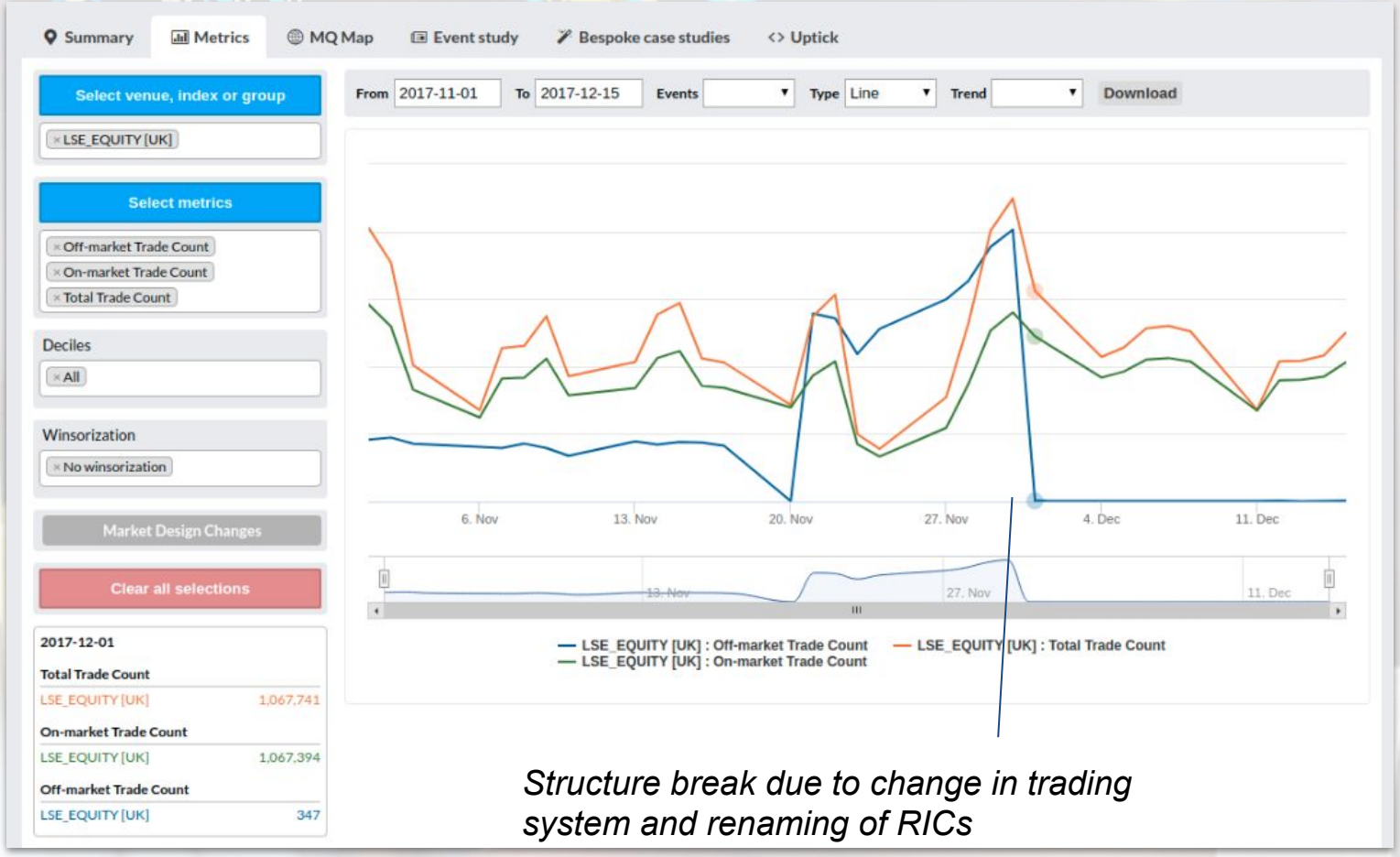


Brexit Warning (maket-wide)

Missing data (jobs ran out of order)

MQD Visual Checks

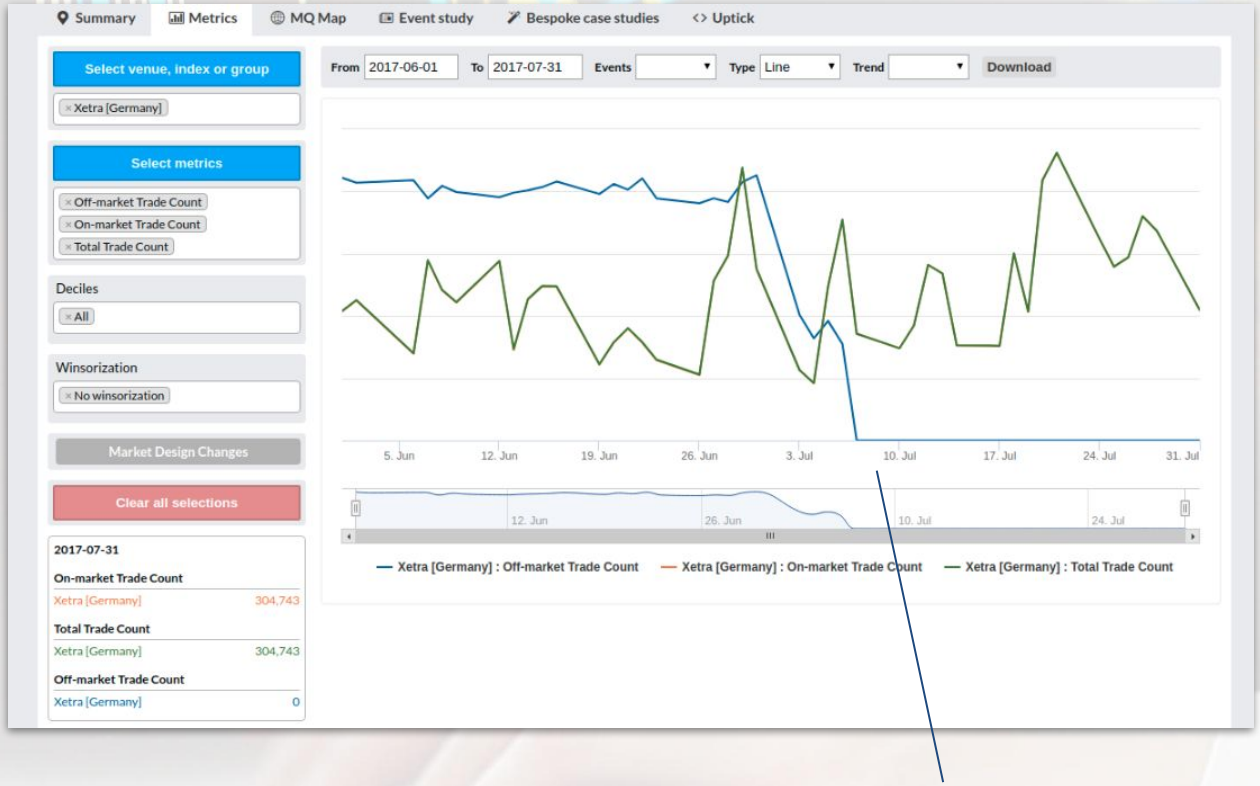
Further examples



Structure break due to change in trading system and renaming of RICs

MQD Visual Checks

Further examples



Another structure break due to change in trading system and removal of qualifiers.

In all cases it was crucial to check the stock exchange website as well as the intraday TRTH output (see next slide)!

MQD Visual Checks

Further examples

TRTH data can be faulty which is usually depicted in MQD as a spike or dip. Here is an example of a data error by TRTH which was fixed in market config:

RIC	Date[L]	Time[L]	Type	Price	Volume	Market VWAP	Bid Price	Bid Size	Ask Price	Ask Size	Qualifiers
OASJ.J	20101213	14:02:09.668050	Quote				1211	3000	1220	10000	
OASJ.J	20101213	14:02:09.871903	Quote				1211		0	0	
OASJ.J	20101213	14:02:09.871903	Trade	1220	10000						" [ACT_FLAG1];Open|High|Low[USER]"
OASJ.J	20101213	14:02:24.983150	Quote				1211		1220	1282156	
OASJ.J	20101213	14:02:25.186176	Quote				1211		0	0	
OASJ.J	20101213	14:02:25.186176	Trade	1220	1282156						" [ACT_FLAG1]"
OASJ.J	20101213	14:02:40.535482	Quote				1211		12200	400	
OASJ.J	20101213	16:50:07.254499	Quote				1211		0	0	
OASJ.J	20101213	19:22:44.023403	Trade								" [PRC_QL_CD]"
OASJ.J	20101213	19:22:46.580569	Trade								Open[USER]

The size of these quotes does not make sense

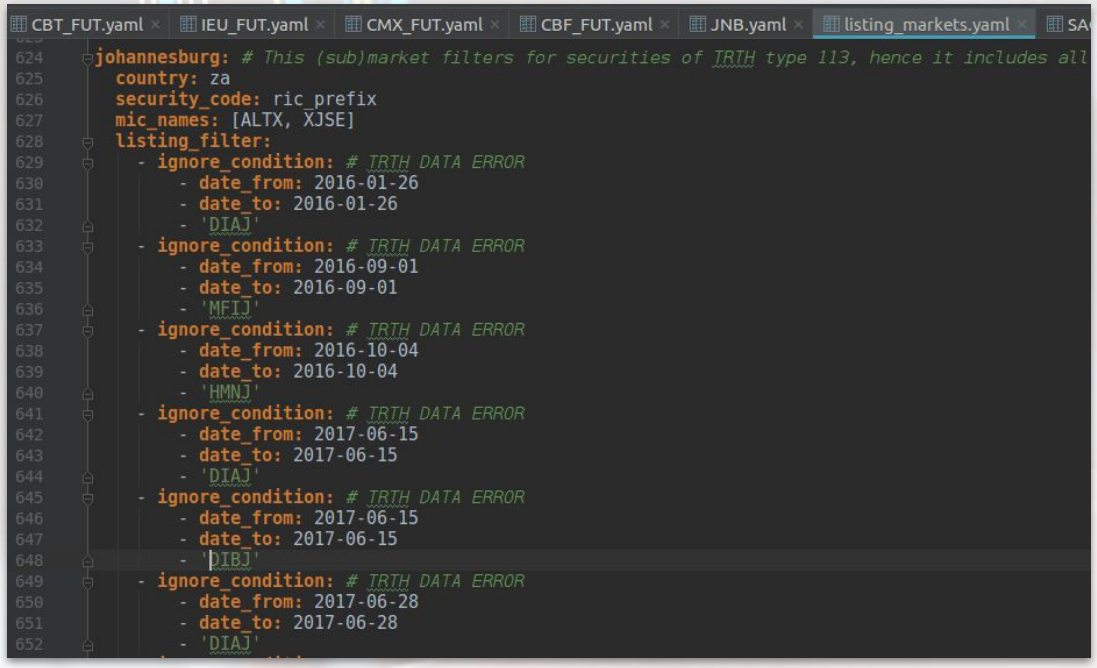
RIC	Date[L]	Time[L]	Type	Price	Volume	Market VWAP	Bid Price	Bid Size	Ask Price	Ask Size	Qualifiers
DIBJ.J	20171006	06:31:24.913800	Quote				0	0	0	0	" [ASK_TONE]; [BID_TONE]"
DIBJ.J	20171006	07:00:00.217633	Quote				450	4500			" [BID_TONE]"
DIBJ.J	20171006	08:46:30.175987	Quote						1030	20000	" [ASK_TONE]"
DIBJ.J	20171006	11:08:33.455716	Quote						0	0	" [ASK_TONE]"
DIBJ.J	20171006	11:10:20.273946	Quote						1030	312909	" [ASK_TONE]"
DIBJ.J	20171006	11:10:20.293267	Trade	1030	312909						" [ACT_FLAG1];44888251116794 1972 [GN_TX20_3]"
DIBJ.J	20171006	11:10:20.293267	Quote						0	0	" [ASK_TONE]"
DIBJ.J	20171006	11:10:20.293292	Trade	1030							Open[USER]
DIBJ.J	20171006	15:17:32.657212	Quote				455	55000			" [BID_TONE]"
DIBJ.J	20171006	15:17:45.916482	Quote						10031	5000	" [ASK_TONE]"
DIBJ.J	20171006	16:50:02.455989	Quote						0	0	" [ASK_TONE]"
DIBJ.J	20171006	16:50:02.455989	Quote				450	4500			" [BID_TONE]"
DIBJ.J	20171006	20:00:53.716746	Quote				450		0		

MQD Visual Checks

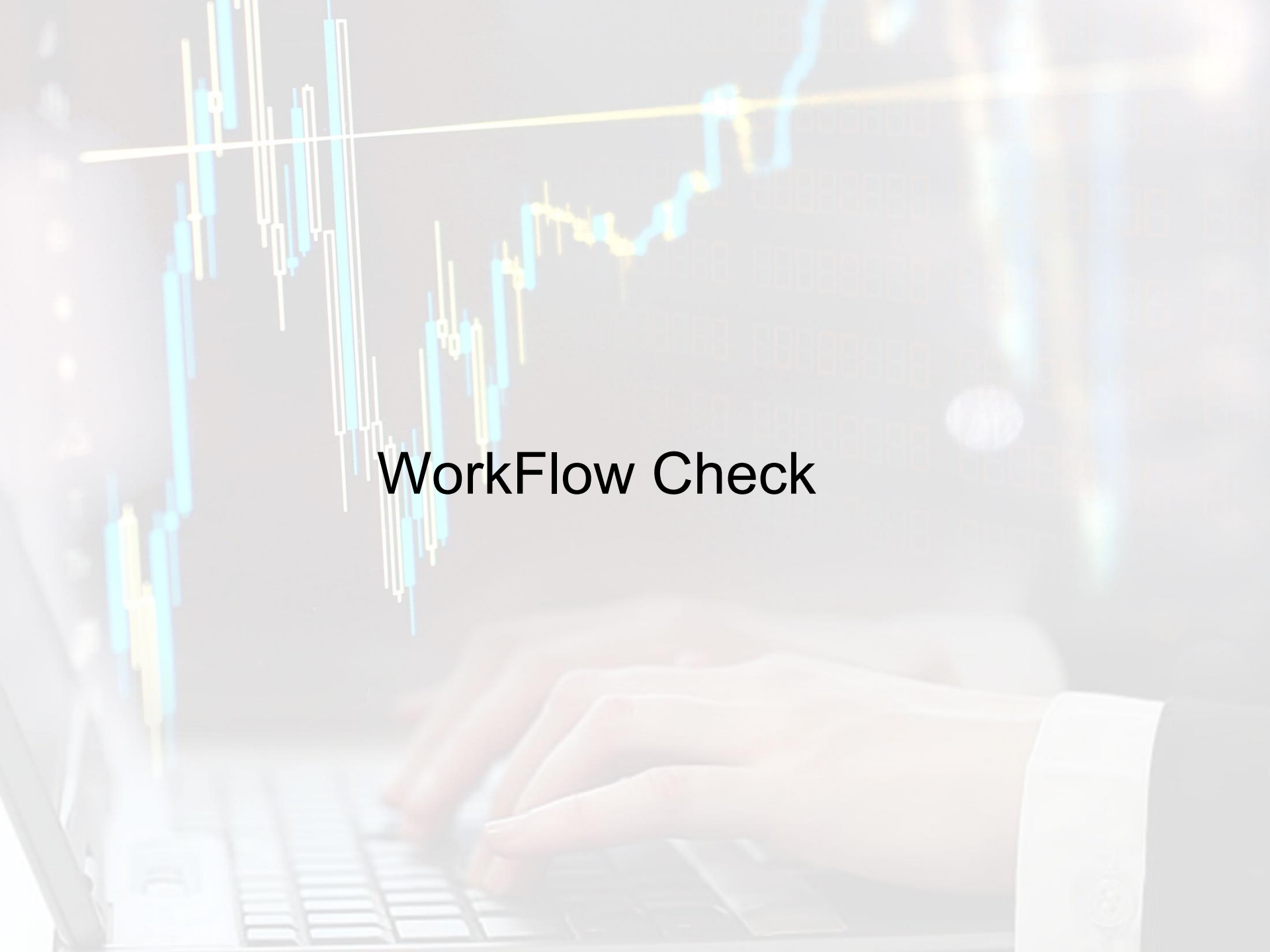
Further examples

TRTH data can be faulty which is usually depicted in MQD as a spike or dip. Here is an example of a data error by TRTH which was fixed in market config:

The solution in this case is to exclude this security for this day.



Careful, if the spike is coming from an event (see Brexit), then we do not exclude or hide it. It belongs to regular trading and market activity! Hiding events is a possibility and needs to be discussed with the teamleads.

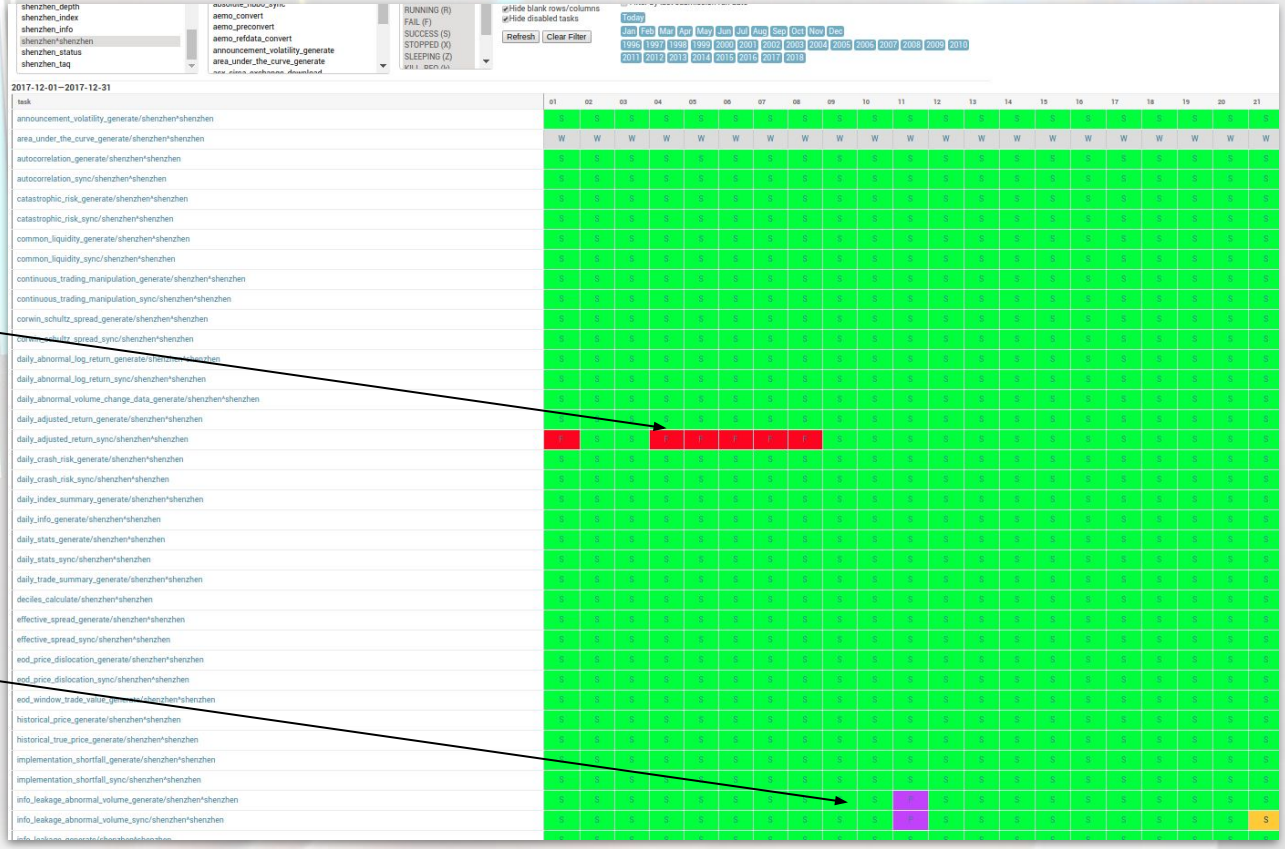

WorkFlow Check

WorkFlow Check

http://WorkFlow.aws.cmcrc.com/

Why did these jobs fail?

Be aware why jobs are pretended in your market!

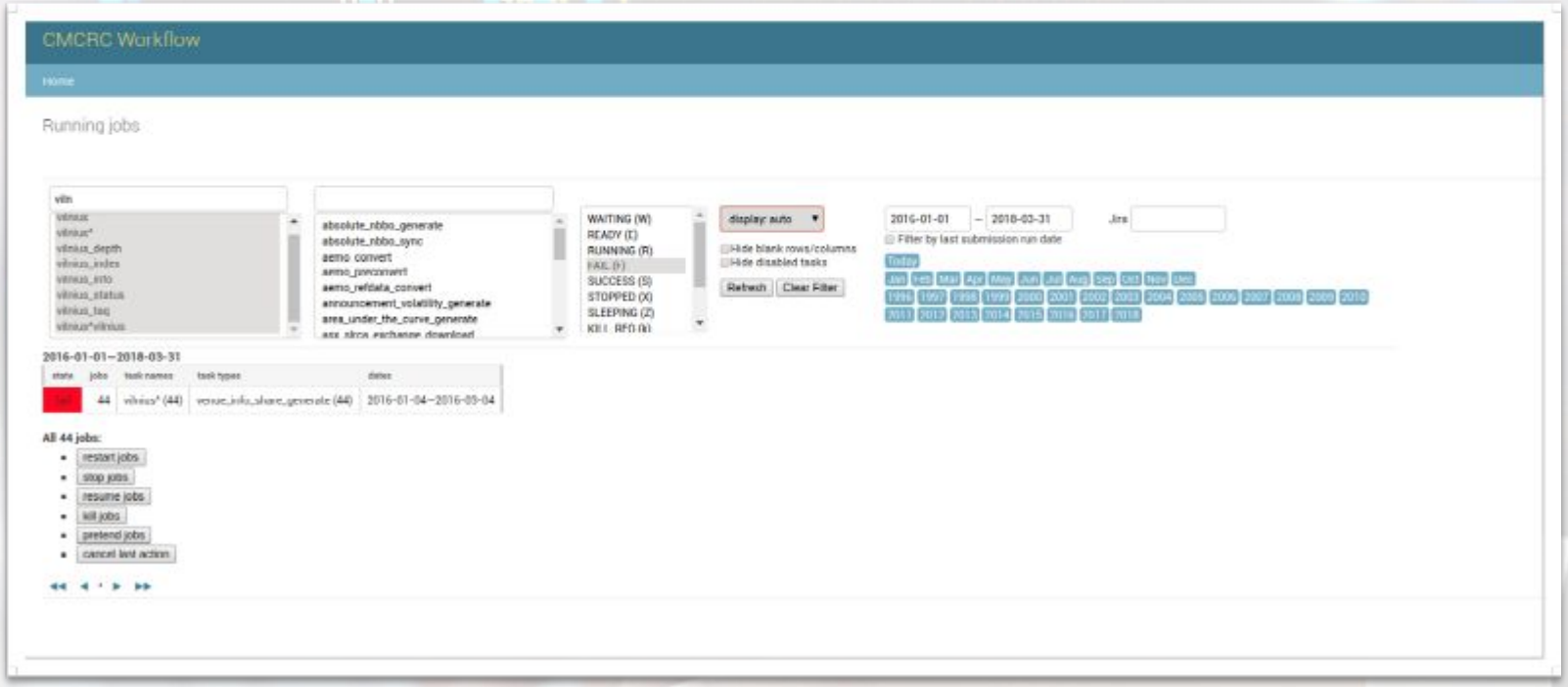


Note: After 100 fails, jobs are deactivated automatically!

WorkFlow Check

http://WorkFlow.aws.cmcrc.com/

Check the failed job summary on a daily basis to get an overview:



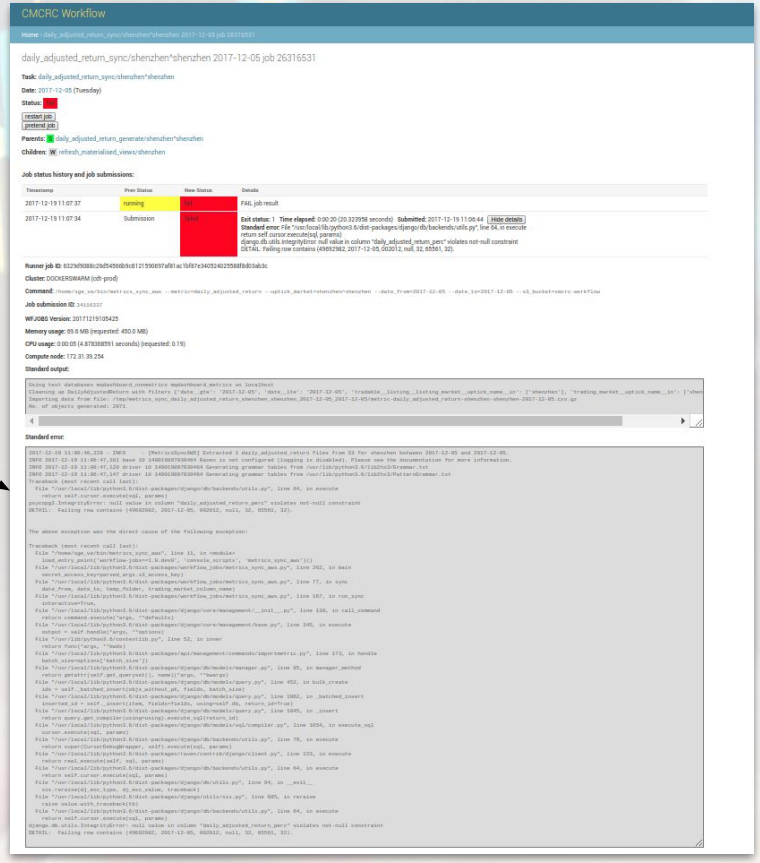
WorkFlow Check

WorkFlow checking

Analyse errors!

The reasons for failing can include:

Codeing bugs
Missing data
Qualifier issues
Currency issues
Dependency issues
Memory/time out issues
.
.
.



Excursion: Opening a JIRA Ticket

After we found an error/spike/structure break or anything that needs our attention, we open a JIRA ticket

A JIRA ticket serves the function of documentation and assignment to the appropriate person

We distinguish between BUS, ETL, and PDEV tickets

- BUS tickets are solved by us – the Business team
- ETL and PDEV tickets are solved by the developers

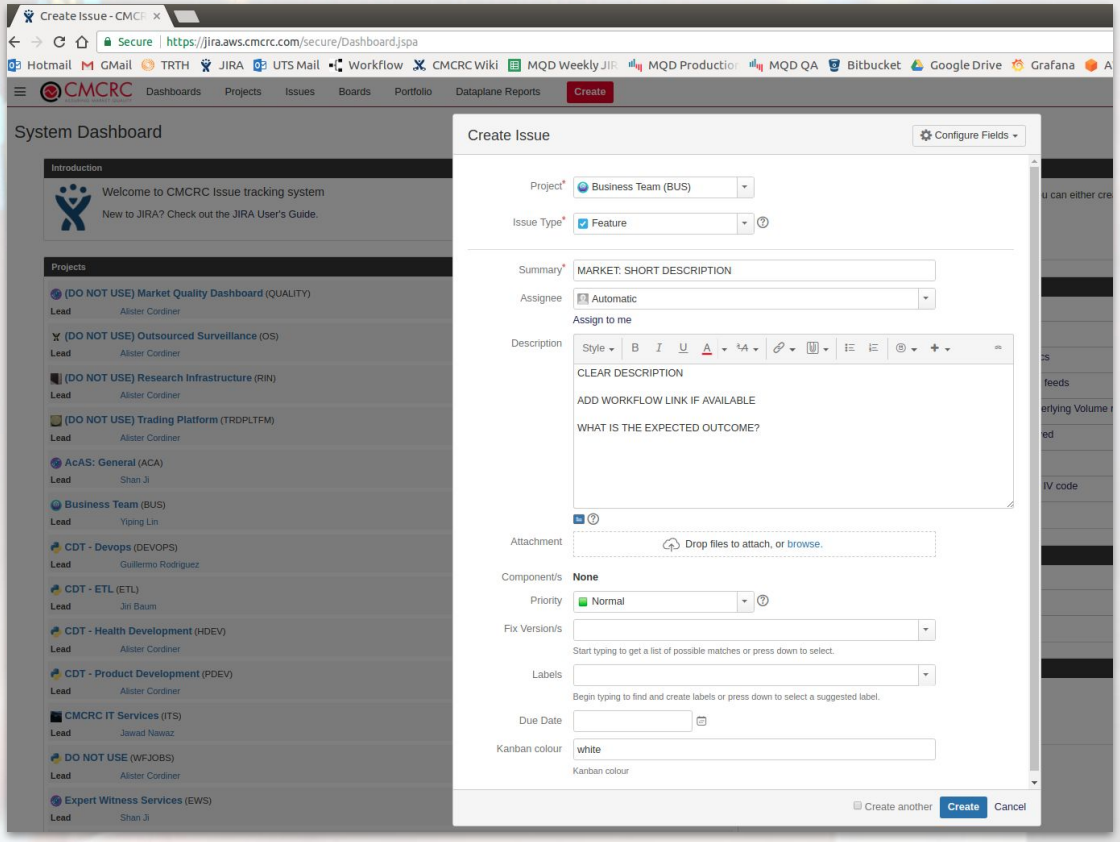
Excursion: Opening a JIRA Ticket

Opening a ticket:

1) Go to jira.aws.cmcrc.com

2) Click on create

3) Follow the steps in the form (live example)



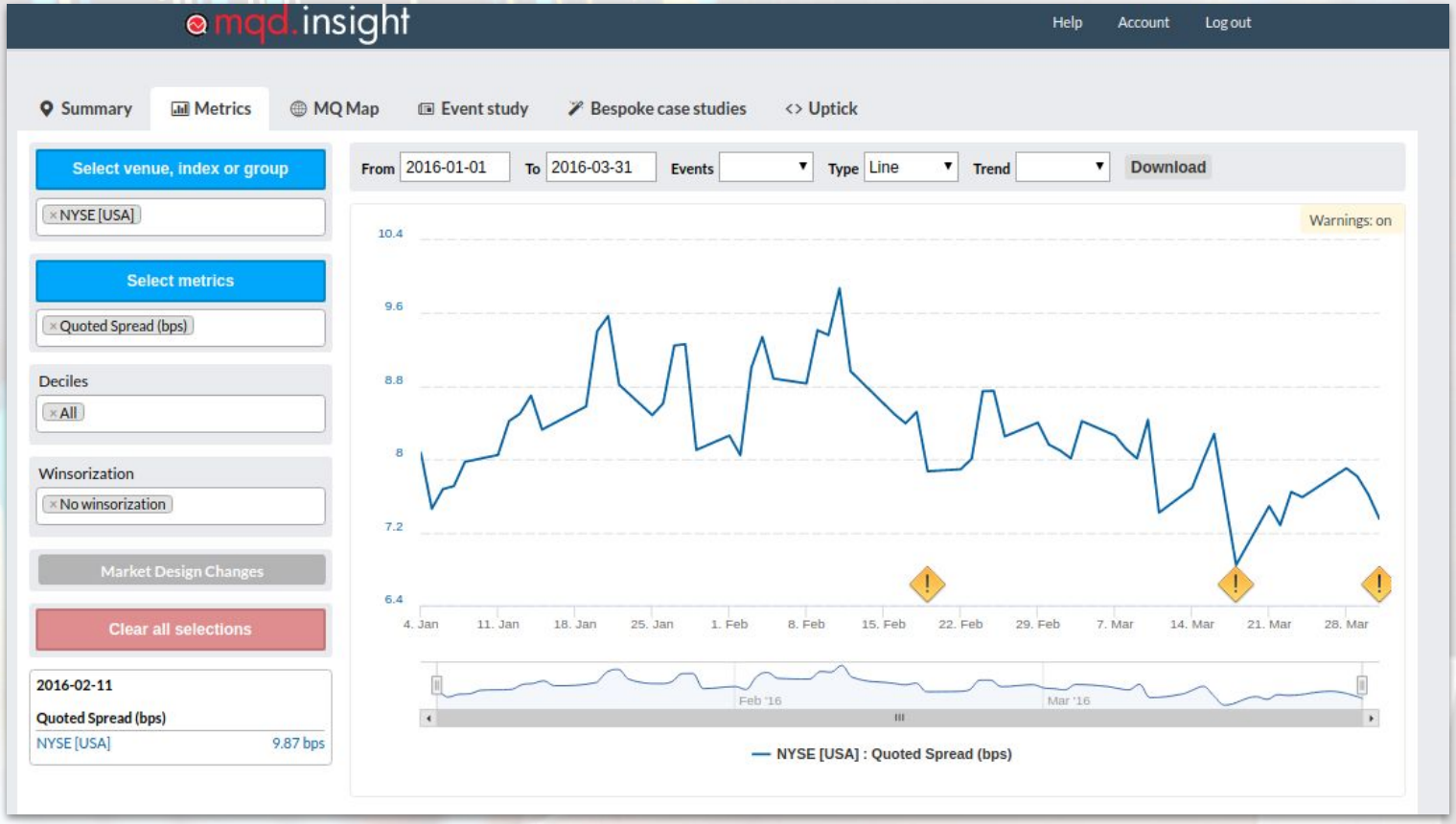
The description must be clear and precisely expressing what the issue is and what is expected as a result!


Random Metric QA

Random Metric QA

QA – one stock

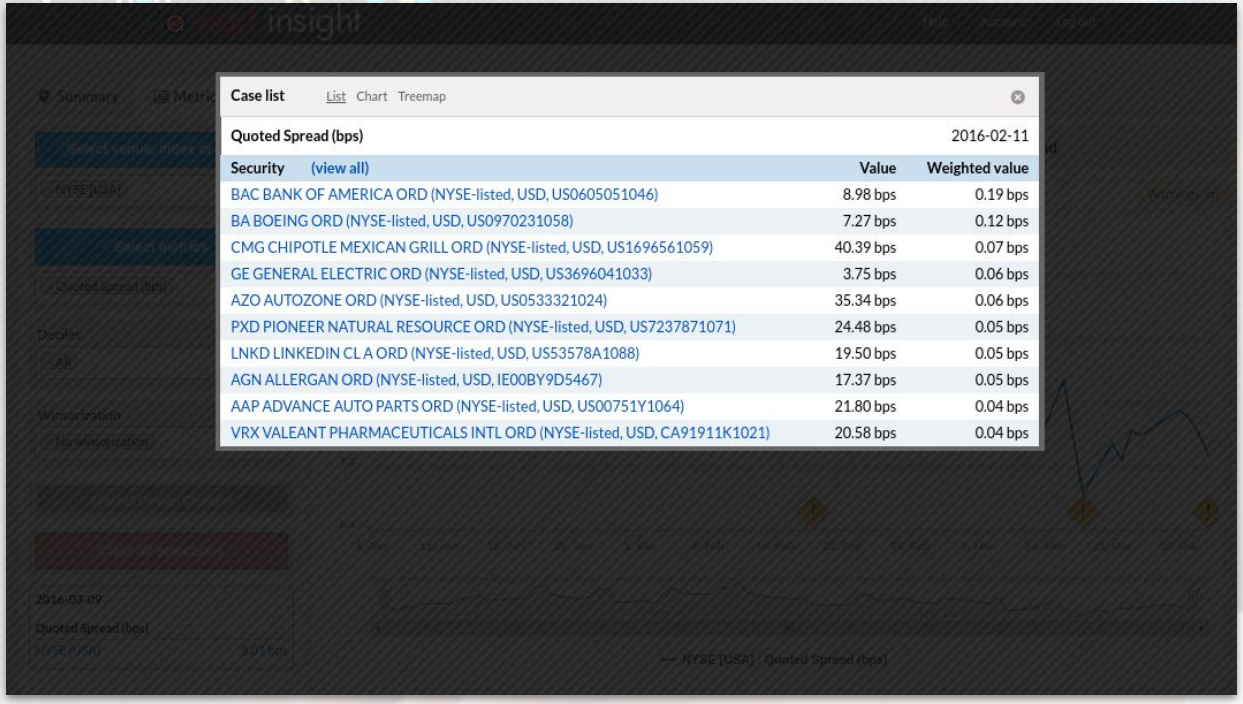
Checking a random day (Quoted Spread) (2016-02-11)



Random Metric QA

QA – one stock

By clicking on the graph, we can see a list of constituents (weighted)

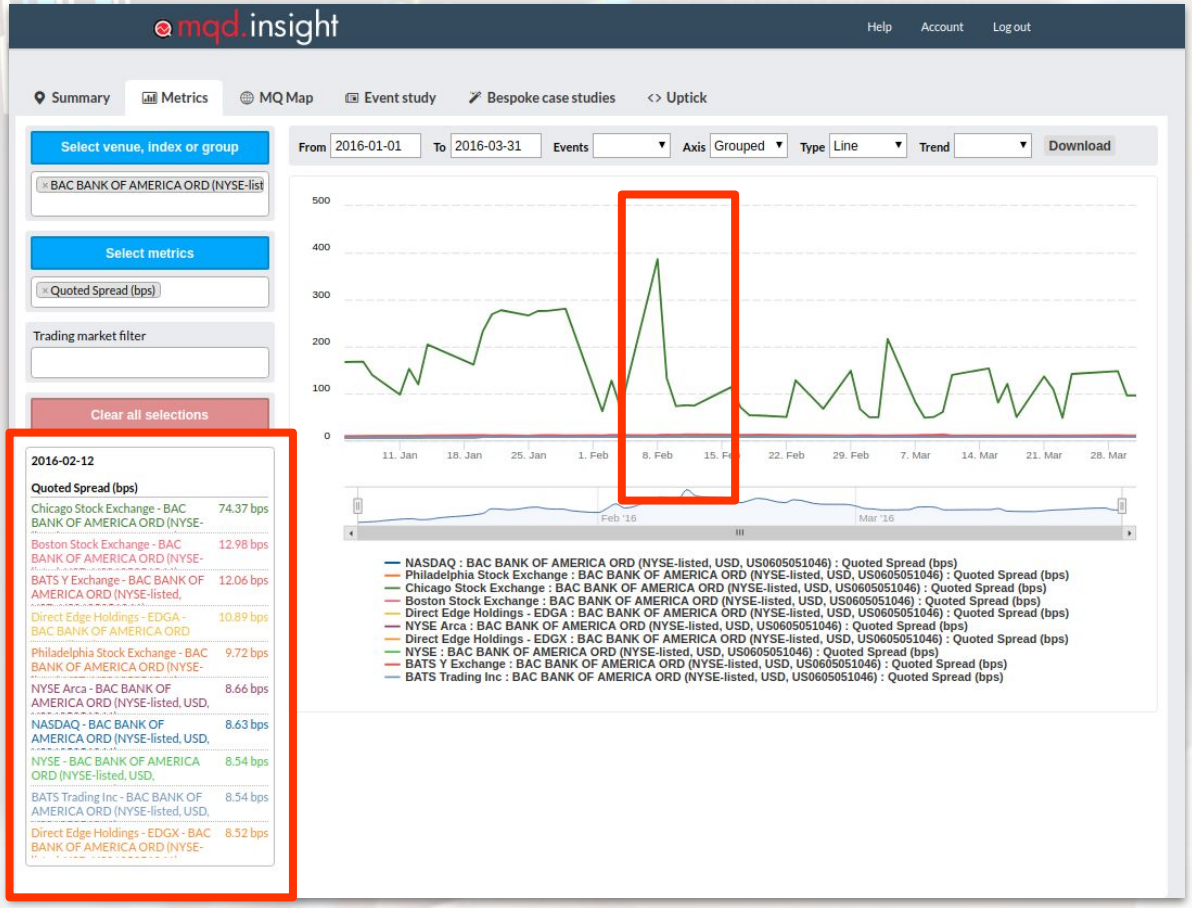


Random Metric QA

QA – one stock

We choose one stock (BAC) and check the quoted spread. The stock is listed on multiple exchanges!

We compare these spreads with

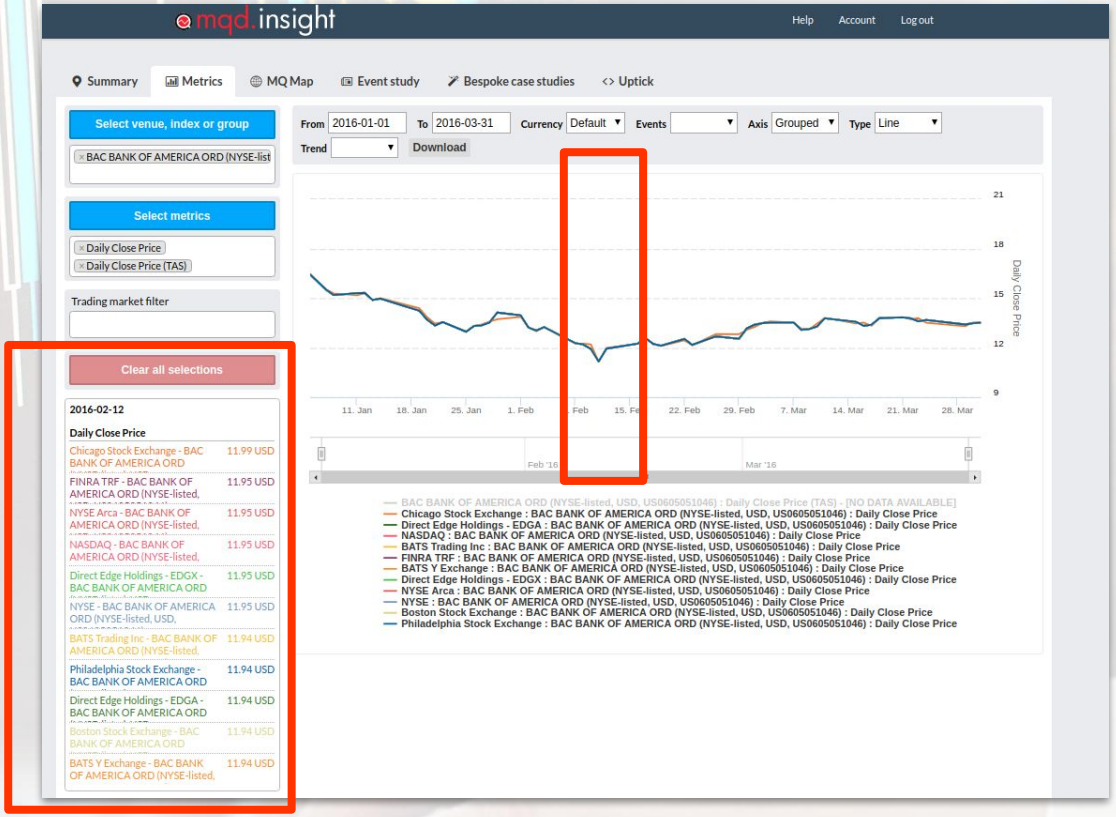


Random Metric QA

QA – one stock

We also want to double check the closing price!

→ Yahoo! Finance

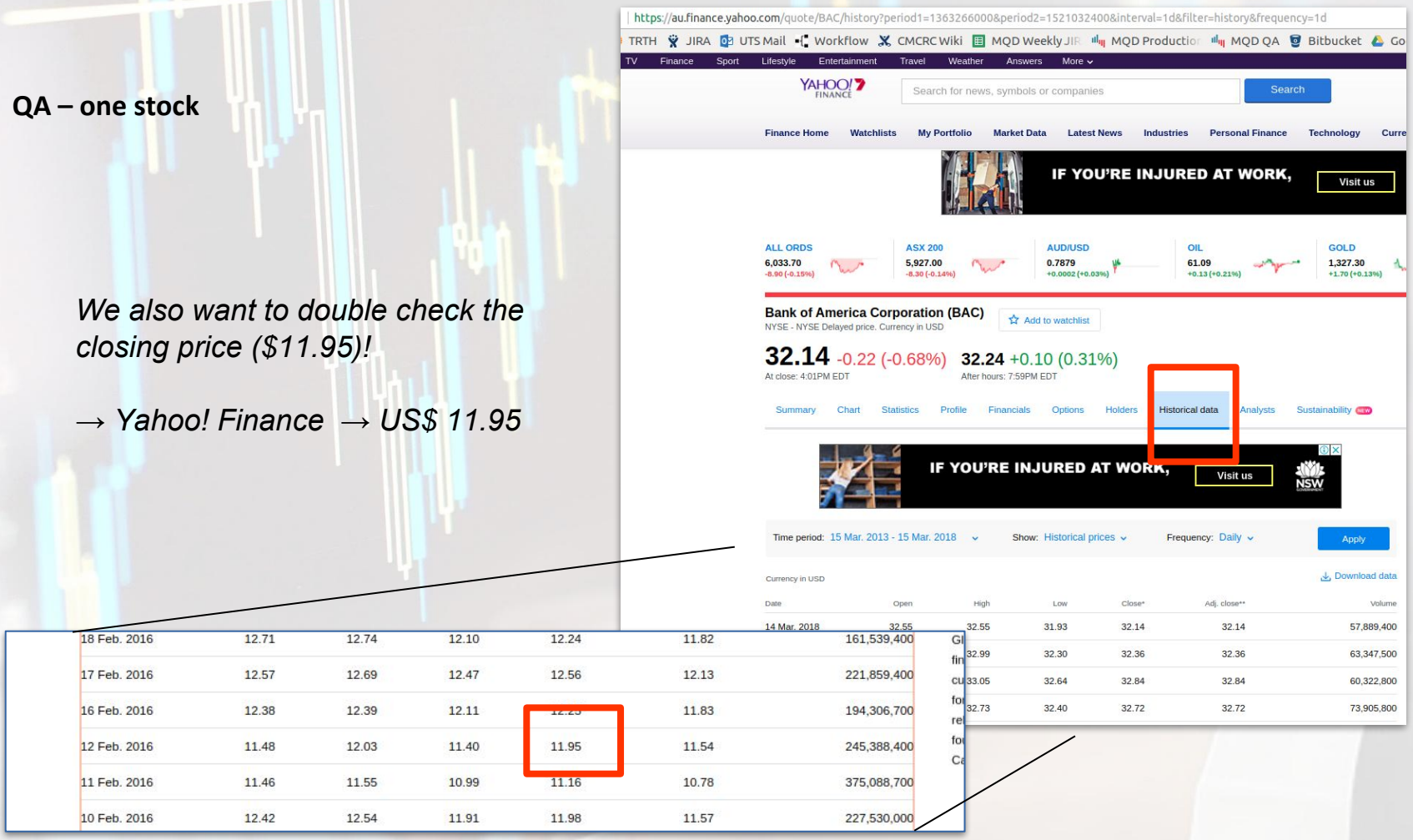


Random Metric QA

QA – one stock

We also want to double check the closing price ($11.95)!

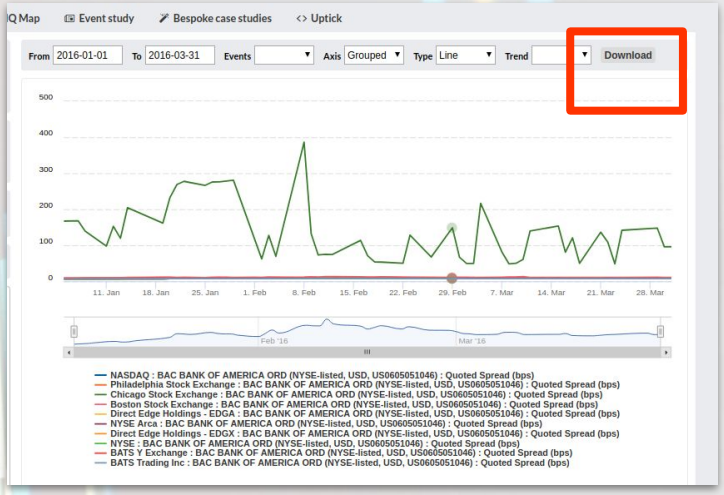
→ Yahoo! Finance → US$ 11.95



Random Metric QA

QA – one stock

Download the data!



We can download the results directly from the MQD website...

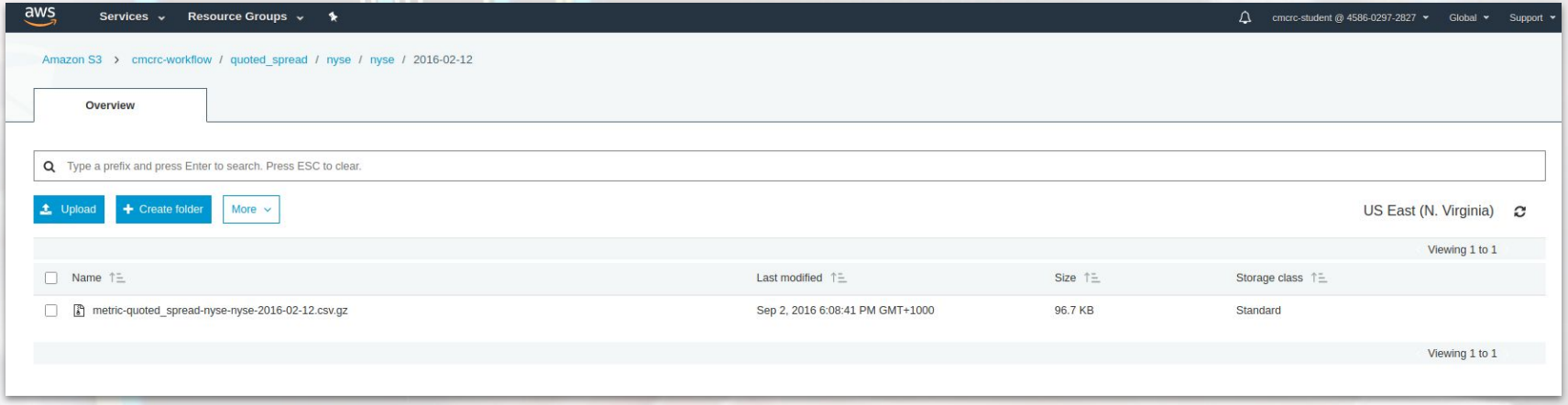
H25 | 8.99 bps

	A	B	C	D	E	F	G	H	I	J	K
1	Date	NASDAQ - BAC	Philadelphia St	Chicago Stock E	Boston Stock Ex	Direct Edge Hol	NYSE Arca - BA	Direct Edge Hol	NYSE - BAC BA	BATS Y Exchan	BATS Trading I
2	2016-01-05	6.42 bps	7.35 bps	167.23 bps	9.45 bps	8.12 bps	6.17 bps	6.11 bps	6.11 bps	8.69 bps	6.15 bps
3	2016-01-07	6.49 bps	8.38 bps	168.26 bps	9.49 bps	8.43 bps	6.49 bps	6.40 bps	6.42 bps	8.81 bps	6.44 bps
4	2016-01-08	6.57 bps	8.43 bps	139.59 bps	10.28 bps	8.30 bps	6.56 bps	6.50 bps	6.51 bps	9.23 bps	6.52 bps
5	2016-01-11	6.72 bps	8.77 bps	97.65 bps	10.13 bps	8.73 bps	6.68 bps	6.62 bps	6.63 bps	9.39 bps	6.64 bps
6	2016-01-12	6.64 bps	8.22 bps	152.67 bps	10.23 bps	8.75 bps	6.63 bps	6.58 bps	6.57 bps	9.37 bps	6.62 bps
7	2016-01-13	6.67 bps	7.79 bps	119.53 bps	10.14 bps	8.51 bps	6.67 bps	6.62 bps	6.62 bps	9.50 bps	6.65 bps
8	2016-01-14	6.78 bps	8.16 bps	204.38 bps	10.64 bps	8.98 bps	6.74 bps	6.71 bps	6.67 bps	9.55 bps	6.74 bps
9	2016-01-19	7.16 bps	8.65 bps	161.61 bps	11.39 bps	9.62 bps	7.13 bps	7.08 bps	7.08 bps	10.60 bps	7.16 bps
10	2016-01-20	7.55 bps	9.33 bps	232.16 bps	11.53 bps	9.99 bps	7.56 bps	7.42 bps	7.43 bps	10.87 bps	7.45 bps
11	2016-01-21	7.53 bps	8.56 bps	268.92 bps	10.76 bps	9.21 bps	7.50 bps	7.41 bps	7.41 bps	10.11 bps	7.43 bps
12	2016-01-22	7.47 bps	8.65 bps	277.58 bps	11.59 bps	9.21 bps	7.48 bps	7.42 bps	7.42 bps	10.03 bps	7.45 bps
13	2016-01-25	7.70 bps	8.56 bps	266.52 bps	10.47 bps	8.93 bps	7.73 bps	7.65 bps	7.65 bps	9.90 bps	7.67 bps

Random Metric QA

QA – one stock

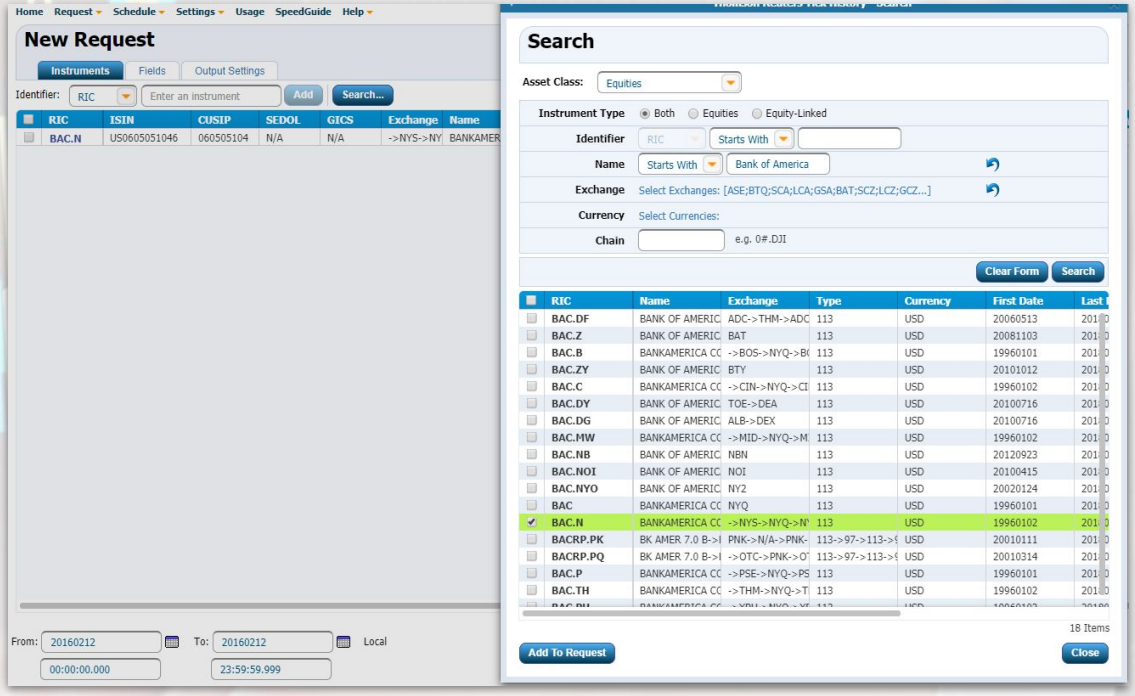
...or get them on a daily basis through Amazon S3 (AWS)



Random Metric QA

QA – one stock

Now we can compare the data with the raw data from TRTH



Random Metric QA

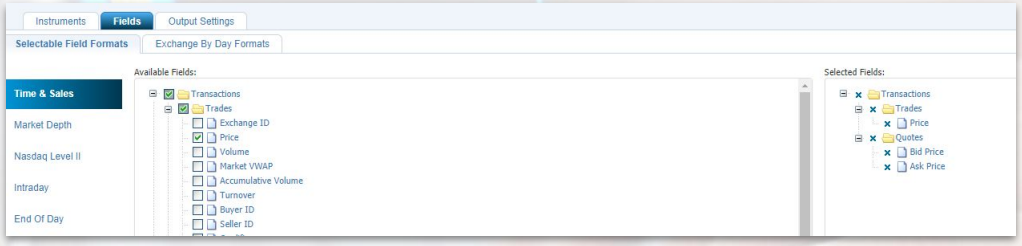
QA – one stock

For Quoted Spread, we just need to download quotes.

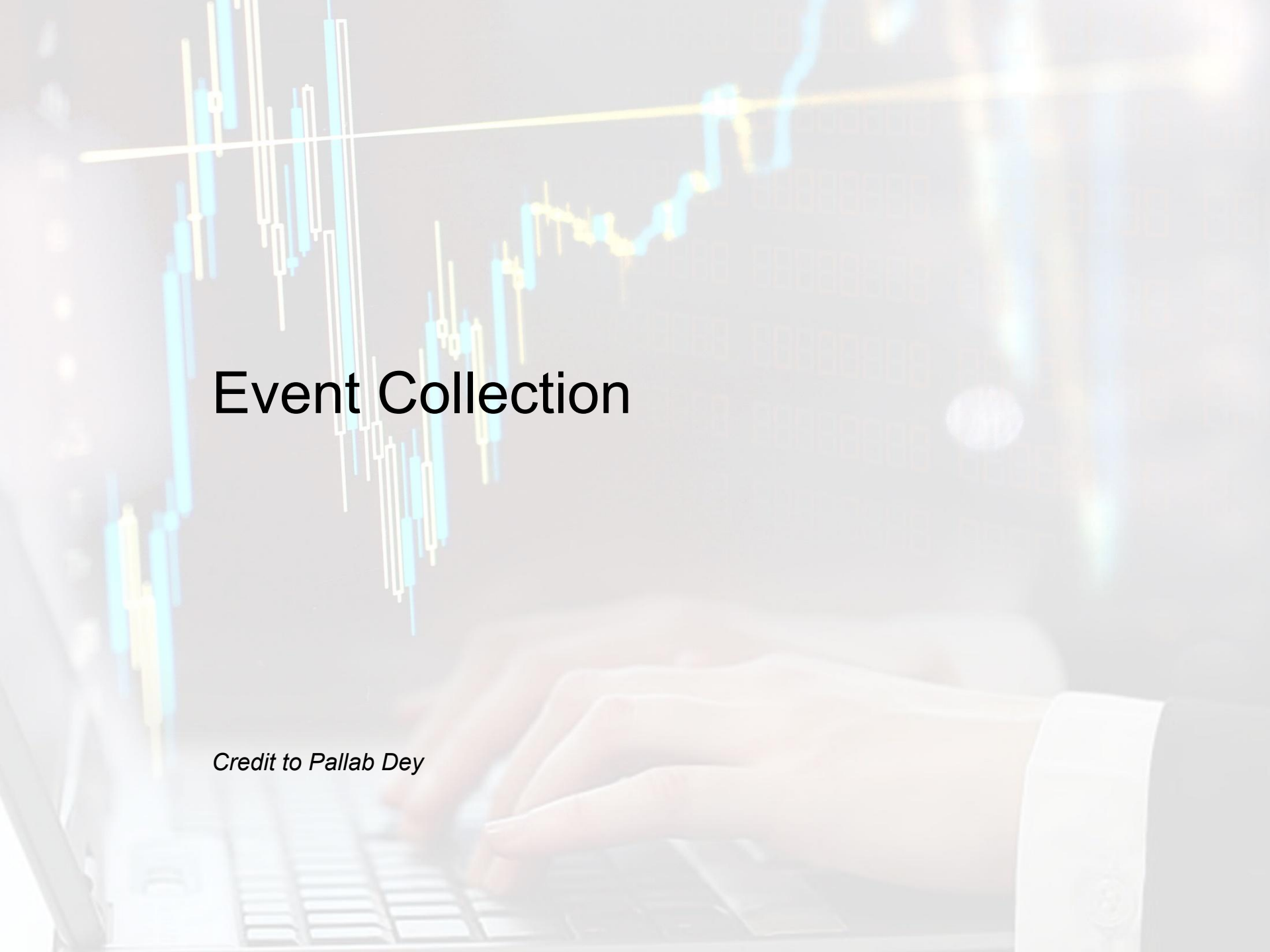
Calculate now the metric by yourself (Excel, Python, …)

The result should be similar to the TRTH value.

Allow for some error as we use precise qualifiers and filtering in MQD.



	A	B	C	D	E	F	G	H	I
1	#RIC	Date[L]	Time[L]	Type	Price	Bid Price	Ask Price		
2	BAC.N	20160212	30:14.7	Trade	11.47				
3	BAC.N	20160212	30:14.7	Quote		11.47	11.48		
4	BAC.N	20160212	30:14.7	Quote		11.47	11.49		
5	BAC.N	20160212	30:14.7	Trade	11.48				
6	BAC.N	20160212	30:14.7	Quote		11.47	11.49		
7	BAC.N	20160212	30:14.7	Quote		11.47	11.49		
8	BAC.N	20160212	30:14.7	Trade	11.49				
9	BAC.N	20160212	30:14.7	Trade	11.47				
10	BAC.N	20160212	30:14.7	Trade	11.49				
11	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
12	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
13	BAC.N	20160212	30:14.7	Trade	11.47				
14	BAC.N	20160212	30:14.7	Trade	11.46				
15	BAC.N	20160212	30:14.7	Trade	11.45				
16	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
17	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
18	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
19	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
20	BAC.N	20160212	30:14.7	Trade	11.45				
21	BAC.N	20160212	30:14.7	Trade	11.49				
22	BAC.N	20160212	30:14.7	Trade	11.49				
23	BAC.N	20160212	30:14.7	Trade	11.49				
24	BAC.N	20160212	30:14.7	Trade	11.49				
25	BAC.N	20160212	30:14.7	Trade	11.49				
26	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
27	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
28	BAC.N	20160212	30:14.7	Quote		11.45	11.49		
29	BAC.N	20160212	30:14.7	Quote		11.48	11.49		
30	BAC.N	20160212	30:14.7	Quote		11.48	11.49		
31	BAC.N	20160212	30:14.7	Quote		11.48	11.49		
32	BAC.N	20160212	30:14.7	Quote		11.48	11.49		
33	BAC.N	20160212	30:14.7	Trade	11.48				



Event Collection

Credit to Pallab Dey

What are Market events

- ‘Exchange and Regulatory related changes that impact market design in Financial Markets. ‘
  - Examples:
    - Tick size changes in Markets – impacts market liquidity/efficiency;
    - Short sell restrictions;
    - New trading venues, instruments types;
    - Technology change – low latency trading engine, co-location;
    - Anti-market manipulation rules.
- Event attributes for MQD:
    - Title, information content, Event categories, Effective date of the event.

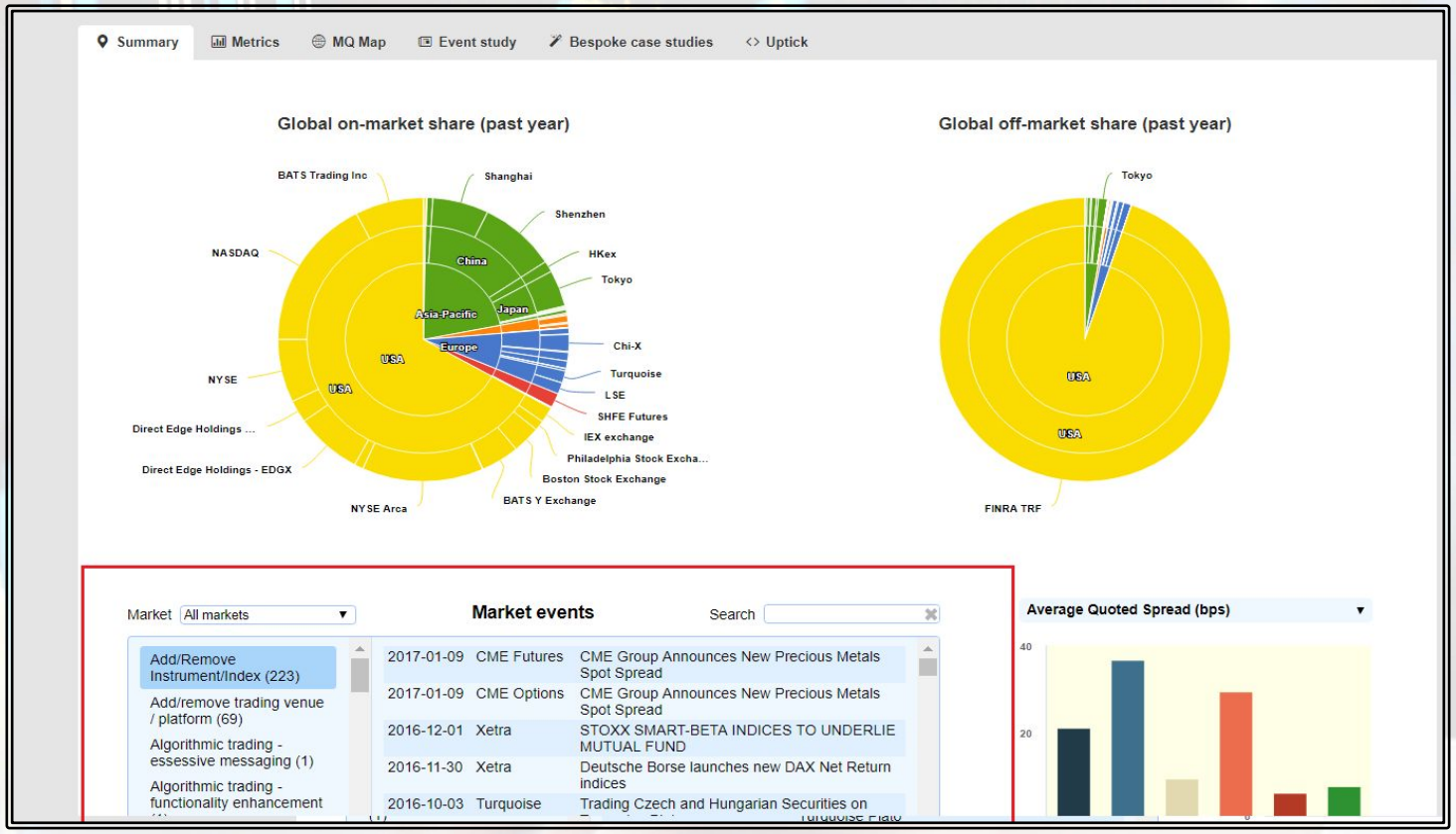
Events in MQD (Metrics)

Important difference: Warning vs Event!

- A warning is marked as an exclamation mark with yellow background and describes data errors and other changes other important issues with the data where the user should be aware of.

- An event is depicted as a number on white background. Events are effects on the market design as described previously.

Events in MQD (MQD Summary)



Steps for Event Collection

- **Identify event source:** exchange news website, regulatory website – For example ASIC (Australia), SEC (US), MAS (Singapore), FCA (UK)
  - Pick up the events that are relevant for MQD

- Collect Event titles, url, date. The events are collected in "**Ongoing Event Collections**" tab in the shared sheet: https://docs.google.com/spreadsheets/d/18CeWkHauTufRg0ksDA5Sqx0kA0Hf5fpkiJ05exLIzMI/edit#gid=1568489002

- **Assign event categories:**
  - TRIPI model (Technology, Regulation, Instrument, Participant, Information)
  - Types of Events – Refer to "Current Event Types" for the list of categories in the shared sheet. The categories configured as dropdown list in the event collection tab.

- Assign a unique id to events: (***ID Format*: market_id + "d"+ date (yyyy/mm/dd) + "n" + number)**. For example, if there are two events on Nasdaq on 2016-09-18, the first one should be 7d20160918n1 and the second one: 7d20160918n2

- Save the event as PDF with "unique id as the file name in following shared location: https://drive.google.com/drive/folders/0BxXihysuXrhAdjBqNElwOW8yaTg

- Capture effective date as event date. Example situations:
  Example1: STOXX events (Announcement date: 01-Aug-2016, Effective Date: 08-Aug-2016. Event date should be: 08-Aug-2016)

PDF File

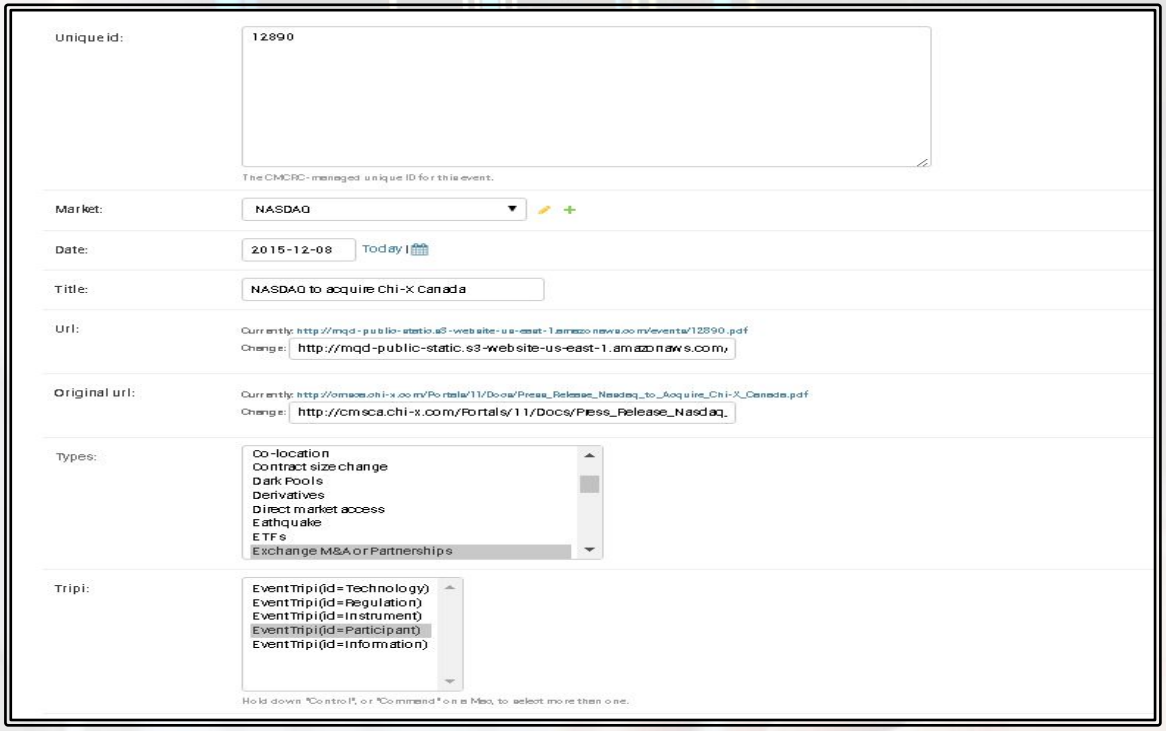
If any announcements causes temporary shocks/spike – this should not be added as events. This needs to be captured as warning (a separate feature in MQD)

Steps for Event Collection (Contd..)

- What is not important for MQD market events? Events that are irrelevant to market microstructure design changes. Just to name few:
  - News about trading statistics in an exchange.
  - Individual securities information.
    - Example: "*SGX Welcomes China Jinjiang Environment to Mainboard*"
  - Trading holidays and half-days (to be captured as warning)
  - Platform version upgrades.
  - Exchange testing:
    - Example: "*ASX Cash Market Services conducts BCP Testing*"
  - Any general news. Example:
    - *"SGX INITIATIVES TO GROW AND DEVELOP THE TR COMMUNITY"*
    - *"Update from Bats Global Markets: New Logo and Naming Conventions"*

- *As a general rule of thumb: Do not ignore if you are not sure whether the particular event is relevant for MQD. This can be further reviewed and filtered out if necessary.*

Interface to add New Events in MQD (MQD Admin / Teamleads)



Adding an event using MQD Admin

- Specify unique id (format: ***market_id + "d"+ date (yyyy/mm/dd) + "n" + number)***
- Specify Market
- Event date
- Title
- URL as the public weblink of the event.
- Original URL – Copy from URL
- Choose Type of event – event category and TRIPI option.
- Click Save

<u>Additional Action :</u>

- Save the event as PDF file using the unique id as file name.
- Create a JIRA and attach the event PDF file (the PDF file needs to be uploaded using back-end script)