BMLL Data Feed

Market State

Date: 2026-04-14
Version: 3.4

1.1. Introduction

BMLL's normalised representation of the trading phase a market is in. Fourteen distinct states cover every venue and mechanism BMLL supports, so cross-venue analysis is directly comparable without exchange-specific logic.

1.2. Dataset schema

Basic Types

Type	Content	Text representation
BOOLEAN	A True or False boolean value.	1,0
INT32	A 32 bit signed integer.	12345
INT64	A 64 bit signed integer.	12345
DOUBLE	IEEE 64-bit floating point value.	123.4567
BYTE_ARRAY	An arbitrarily long byte array.	"VODAFONE GROUP"
NUMBER	A number with defined precision (total digits allowed) and scale (number of digits allowed to the right of the decimal point). Unless otherwise specified, the default is precision 38 and scale 0.	12345
VARCHAR	Variable length string up to 500 characters. May contain quotes and special characters. All text will be utf8-encoded unless indicated differently. Ticker and ISO codes should be ASCII compatible.	"VODAFONE GROUP"
TIMESTAMP_NTZ	Nanoseconds since 1970 in UTC. Timestamp will be represented in ISO format in text format.	20220302-11:23.57.12345789
FLOAT	IEEE 64-bit floating point value.	123.4567
DATE	A calendar date consisting of a year, month, and day, without including any time or time zone components.	2025-01-01

T: +44 (0)20 3828 9000
E: info@bmlltech.com
W: www.bmlltech.com
Copyright © 2026 BMLL Technologies. All rights reserved.

1.3. Schema

One row corresponds to a market state change event for a specific instrument.

Field Name (Snowflake / Databricks)	Field Name (Parquet)	Type (Snowflake / Databricks)	Type (Parquet)	Description
TRADE_DATE	TradeDate	DATE	INT32	Date associated with the dataset row. Product-level smell records document any legacy product where this date is derived from publication time rather than the venue trade timestamp.
TIMESTAMP_NANOSECONDS	TimestampNanoseconds	NUMBER	INT64	This column contains a representation of the time of the update as the number of nanoseconds since 1970-01-01 00:00:00.0 UTC.
MARKET_STATE	MarketState	VARCHAR	BYTE_ARRAY	Standardised market state information. See below for description of possible market states.
MIC	MIC	VARCHAR	BYTE_ARRAY	The Market Identifier Code (MIC) is used to uniquely identify the exchange which is reporting market activity.
EXCHANGE_MARKET_STATE	ExchangeMarketState	VARCHAR	BYTE_ARRAY	Flags related to market state as implemented by exchange. This field contains an exchange specific values that represent the latest exchange-specific flags associated with a change of market state on the main order lit order book.

Field Name (Snowflake / Databricks)	Field Name (Parquet)	Type (Snowflake / Databricks)	Type (Parquet)	Description
LISTING_ID	ListingId	NUMBER	INT64	The BMLL ListingId is used to identify a specific instrument on a market venue (trading venue or reporting venue). BMLL listings identify uniquely every ticker on the venue and remain attached to the instrument even when a ticker change occurs — though changes of primary instrument identifier (such as ISIN changes) can cause a new listing_id to be issued.
INSTRUMENT_ID	InstrumentId	NUMBER	INT64	The BMLL identifier of the BMLL instrument. BMLL instruments are designed to connect together listings of a designated underlying asset when they are issued with similar condition (i.e. same quoting currency, geographical, legislation region, primary reference venue).
EXCHANGE_SEQUENCE_NO	ExchangeSequenceNo	NUMBER	INT64	The sequence number as provided by the exchange. Note that some exchange may not provide a sequence number in this case the ExchangeSequenceNo will always be 0.

Field Name (Snowflake / Databricks)	Field Name (Parquet)	Type (Snowflake / Databricks)	Type (Parquet)	Description
BMLL_SEQUENCE_NO	BMLLSequenceNo	NUMBER	INT64	The sequence no of the trade in the upstream feed as inferred by BMLL - this sequence number will respect the order of messages in upstream protocol and will be populated even if the upstream venue does not provide a sequence number - where the source of the feed is indicated by BMLLSequenceSource.
BMLL_SEQUENCE_SOURCE	BMLLSequenceSource	NUMBER	INT64	An integer used to identify the upstream feed from which the sequence numbers are generated. This is designed to avoid the creation of non-existent order between events that come from different message streams

1.4. Market states

BMLL normalises each venue's phase messages into a consistent vocabulary. Canonical MarketState enumeration.

Code	Description
AUCTION_ON_DEMAND	An auction that occurs concurrently to the main trading session
CLOSED	The exchange is closed
CLOSING_AUCTION	Market in Closing Auction
CONDITIONAL	The trade is the result of a conditional trading mechanism

1.5. Format and File Structure

Data is delivered as a single file per market per date, as a parquet file.

- **{YYYY}:** Year
- **{MM}:** Month
- **{DD}:** Date
- **{Product}:** The product you are subscribing
  - **Market State:** MARKETSTATE
- **{Region}:** The region of the venue such as APAC, AMERICAS, and EMEA
- **{MIC}:** Market Identifier Code (MIC) is a unique four-character code that identifies stock markets and exchanges for trading and referencing computer systems

File Structure:

{YYYY}/{MM}/{DD}/MARKETSTATE/{Region}/market-state-{MIC}-{YYYYMMDD}.parquet

Example:

2026/04/14/MARKETSTATE/EMEA/market-state-XLON-20260414.parquet