



US Futures Trade and Quote Minute Bar Guide

version 2.2 (Jul 2021)



algoseek | the market data company

We provide research market data for machine learning and quantitative trading



CONTACT US

We are here to help you do great things with our market and reference data. For questions, feedback, and other concerns, you may reach our team of experts using the following contact information:

algoseek customer support

 support@algoseek.com

 (+1) 646 583 1832

algoseek sales

 sales@algoseek.com

 (+1) 646 583 1832



TABLE OF CONTENTS

INTRODUCTION	4
MINUTE BARS CALCULATIONS	4
DATA ORGANIZATION AND FILE FORMAT	5



INTRODUCTION

algoseek's Futures, Spreads, and Futures Options Trade and Quote Minute Bar datasets contain high-quality intraday information-rich one-minute bar data for all securities listed on US exchanges. The data includes event-based bars designed for quantitative trading, backtesting, machine learning, and other advanced applications.

The contract specifications such as underlying asset, the quantity of the asset or contract size, expiration months, delivery location, and the delivery date upon execution can be found on the US futures products list, which is available on the following link: <http://www.cmegroup.com/trading/products/>.

MINUTE BARS CALCULATIONS

Event-based Bars

The volume of futures varies from contract to contract. Some contracts, for example, ES front month, have multiple events per millisecond, and some contracts don't have any type of event (i.e. bid, ask, or trade) for 24 hours.

algoseek's futures bars are event-based, which means a bar is only created when there is at least one event during a bar period. The three types of events (bid, ask, and trade) don't always happen together during a bar period. For example, there could be three bids, but neither trades nor asks for a minute.

Implied Quotes and Calculated Trade Prices

There are Implied Bid and Ask Quotes for some futures based on spread relationships (e.g., Crude Oil CL futures). However, these are ignored when building the bars. Only real bid and ask quotes are included.

Some futures have calculated prices that are not based on actual trades, and these are not included in bars. Only trades of the actual futures are included.

Daily Halts and Breaks

If the specified futures contract has a daily break, you may see bars with the bid or ask events that the exchange published during the break period. For example, Silver Futures (SI) has a daily hour break from 4 pm to 5 pm Central Time.

These bid or ask events are published by the exchange to indicate the expected price after the daily halt.



Minimum Quotes and Trades

A bar file is only created if there is at least one of the following conditions apply within 24 hours:

- 3 or more quotes
- 1 or more trades

If there is only one type of event during the bar period, the event fields will have the same values. For example:

`OpenBidPrice = HighBidPrice = LowBidPrice = CloseBidPrice`

`OpenBidSize = HighBidSize = LowBidSize = CloseBidSize`

`OpenBidTime = HighBidTime = LowBidTime = CloseBidTime`

Bars are created only if there is a trade or quote event during the minute. For illiquid futures with no trades or quotes on a given trade date, there will be no bar data file. Some very illiquid futures may even not have any trades or quotes for weeks at a time.

DATA ORGANIZATION AND FILE FORMAT

algoseek provides Futures market data in plain-text CSV files. The first row of the CSV file is a fixed header, and then rows of data corresponding to individual events. By default, data is organized into one file per symbol per trading day. For example, all events for ticker ESH0 on Jan 28, 2020, are stored in one CSV file.

Due to the large data size, CSV files are gzip-compressed (having a csv.gz extension) with a compression ratio of about 8:1.

Table 1 below provides the name, base event, default value, brief description, and data type for each data field (column) in the Futures TAQ Minute Bar CSV file. Table column “Missing” indicates a default behavior in case the data field value is not present or cannot be calculated. The column value “Never” means that the data field value is always present.

Table 1: CSV File Fields Schema

Field	Base Event	Type (Format)	Missing	Description
Ticker	-	string	Never	Symbol name for a specific contract
UTCDate	-	string (yyyymmdd)	Never	UTC trade date



UTCTimeBarStart	-	string (time)	Never	For minute bars, the format is HH:MM. For second bars, it is HH:MM:SS.
LocalDate	-	string (yyyymmdd)	Never	Local trade date based on local Chicago time.
LocalTimeBarStart	-	string (time)	Never	Local time in Chicago including daylight saving time changes
CallPut	-	string	Never	Option type (Call or Put) displayed as "C" or "P". Note: available only in Future Options dataset
Strike	-	integer	Never	A fixed price for buying or selling underlying outright future contract. Note: available only in Future Options dataset
Month	-	string	Never	A code for the expiration month for the option contract (as a single letter). Note: available only in Future Options dataset
ExpirationYear	-	integer	Never	The expiration year of the option contract. Note: available only in Future Options dataset
OpenBidTime	Quote	string (timestamp)	Blank	Time of first bid
OpenBidPrice	Quote	decimal	Blank	Opening bid price
OpenBidSize	Quote	integer	Blank	Opening bid Size
OpenAskTime	Quote	string (timestamp)	Blank	Time of first ask
OpenAskPrice	Quote	decimal	Blank	Opening ask price
OpenAskSize	Quote	integer	Blank	Opening ask size
OpenTradeTime	Trade	string (timestamp)	Blank	Time of first trade
OpenTradePrice	Trade	decimal	Blank	Price of first trade
OpenTradeSize	Trade	integer	Blank	Number of contracts of first trade
HighBidTime	Quote	string (timestamp)	Blank	Time of highest bid price
HighBidPrice	Quote	decimal	Blank	Highest bid price



HighBidSize	Quote	integer	Blank	Bid size of highest bid price
HighAskTime	Quote	string (timestamp)	Blank	Time of highest ask price
HighAskPrice	Quote	decimal	Blank	Highest ask price
HighAskSize	Quote	integer	Blank	High ask size
HighTradeTime	Trade	string (timestamp)	Blank	Time of highest trade
HighTradePrice	Trade	decimal	Blank	Price of highest trade
HighTradeSize	Trade	integer	Blank	Number of contracts of highest trade
LowBidTime	Quote	string (timestamp)	Blank	Time of lowest bid
LowBidPrice	Quote	decimal	Blank	Lowest bid price
LowBidSize	Quote	integer	Blank	Bid Size of lowest bid price
LowAskTime	Quote	string (timestamp)	Blank	Time of lowest ask price
LowAskPrice	Quote	decimal	Blank	Lowest ask price
LowAskSize	Quote	integer	Blank	Lowest ask size
LowTradeTime	Trade	string (timestamp)	Blank	Time of lowest trade
LowTradePrice	Trade	decimal	Blank	Price of lowest trade
LowTradeSize	Trade	integer	Blank	Number of contracts of lowest trade
CloseBidTime	Quote	string (timestamp)	Blank	Time of last bid
CloseBidPrice	Quote	decimal	Blank	Price of last bid in bar
CloseBidSize	Quote	integer	Blank	Size of Last bid in bar
CloseAskTime	Quote	string (timestamp)	Blank	Time of last ask
CloseAskPrice	Quote	decimal	Blank	Price of last ask in bar
CloseAskSize	Quote	integer	Blank	Size of last ask in bar
CloseTradeTime	Trade	string (timestamp)	Blank	Time of last trade
CloseTradePrice	Trade	decimal	Blank	Price of last trade



CloseTradeSize	Trade	integer	Blank	Number of contracts of last trade
MinSpread	Quote	decimal	Blank	Minimum Bid-Ask spread size: this requires both a bid and ask during the minute. If the spread is negative, then it is 0.
MaxSpread	Quote	decimal	Blank	Maximum Bid-Ask spread size: This requires both a bid and ask during the minute.
VolumeWeightPrice	Trade	decimal	Blank	Trade Volume weighted average price:
TotalRegularQuotes	Quote	integer	Blank	Number of Regular Bid/Ask quotes. This remains blank if there are no bid or ask events. Implied quotes are ignored for bar creation.
BuyAggressorTrades	Trade	integer	0	Number of trade where type is "Trade Aggressor at Buy"
SellAggressorTrades	Trade	integer	0	Number of trade where type is "Trade Aggressor at Sell"
NoAggressorTrades	Trade	integer	0	Number of trades where exchanges do not publish the Trade Aggressor Side
BuyAggressorQuantity	Trade	integer	0	Sum of trade contracts where type is "Trade Aggressor at Buy"
SellAggressorQuantity	Trade	integer	0	Sum of trade contracts where type is "Trade Aggressor at Sell"
NoAggressorQuantity	Trade	integer	0	Sum of trade contracts where exchanges do not publish the Trade Aggressor Side
Volume	Trade	integer	0	Total number of contracts traded
TotalTrades	Trade	integer	0	Total number of trades

Time Bar Start Format: One-second bar 13:03:02 is from a time greater than 13:03:01 to 13:03:02. One-minute bar 11:04 is from time greater than 11:04 to less than 11:05.

VolumeWeightPrice: the volume-weighted price is calculated as a sum of all individual trade volumes divided by the total number of shares traded

$$\frac{((Trade1_Shares \times Trade1_Price) + (Trade2_Shares \times Trade2_Price) + \dots)}{TotalShares}$$



Timestamps

The timestamps are all in Coordinated Universal Time (UTC). However, each bar also includes fields “LocalDate” and “LocalTimeBarStart” in Chicago local time, which is in the Central Standard Time (CST) zone.

For example, an algoseek Futures timestamp of 181514415 is 18:15:14.415 UTC. During Daylight Saving Time this converts to 13:15:14.415 CST.

Given that UTC varies by five or six hours from Chicago local time, the “LocalDate” field can be different from the “UTCDate” field. Chicago Winter Time is UTC –6 hours, and Summer Daylight Saving Time (DST) is UTC –5 hours.

For example, UTCTimeBarStart time 01:30 (1:30 a.m. UTC) with UTC date 2017-08-15 will have LocalDate 2017-08-14, and LocalTimeBarStart 20:30 (8:30 p.m. CST) due to the 5-hour difference during Daylight Saving Time.

Note: Excel will automatically try to convert millisecond timestamps into Excel time and will fail. Please make sure to convert the millisecond timestamps to text when importing.