



US Equity Daily Trade and Quote Guide

version 1.0 (Jan 2022)



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	5
DATA SOURCE	5
FINRA TRF AND ODD LOT TRADES	6
OFFICIAL OPEN/CLOSE IDENTIFICATION	6
Daily BAR CALCULATIONS	7
EXCHANGE MARKET HOURS	8
DATA ORGANIZATION AND FILE FORMAT	9
APPENDIX A. FREQUENTLY ASKED QUESTIONS	16
APPENDIX B. BAR CALCULATIONS FROM TRADE AND QUOTE EVENTS	18



INTRODUCTION

AlgoSeek Daily Trade and Quote (TAQ) data is built from top-of-book intraday quotes and trades for all listed stocks, ETNs, ETFs, ADRs, and funds from 16+ US exchanges and marketplaces.

Each US Equity has a Primary Listing Exchange considered the official Open/Close auction cross for the Equity even though Open/Close auction crosses may take place on other exchanges. This dataset includes each Equity's Open/Close based on the primary exchange on that trade date. The Primary Exchange for a stock occasionally changes, for example, when ORCL moved to NYSE from NASDAQ in July 2013.

With more than 60 data fields, algoseek Daily TAQ dataset is the most comprehensive and detailed Daily bar product in the financial industry. It's designed for quantitative trading, backtesting, machine learning, and other advanced applications.

Data files are in CSV (Comma-Separated Values) format. An individual CSV file is created for all symbols per trading day

DATA SOURCE

algoseek Daily TAQ dataset is built from "as-is" tick data collected from live SIP feed algoseek's co-located ticker plant servers in Equinix NY2 and NY4 data centers, connected with 10Gb fiber for low latency.

The Securities Information Processor (SIP) includes Tape A and Tape B covered by the Consolidated Tape Association (CTA) plan and Tape C covered by the Unlisted Trading Privileges (UTP) plan. The SIP links the US markets by processing and consolidating all protected bid/ask quotes and trades from every trading venue into a single and easily consumable data feed.

The SIP disseminates and calculates critical regulatory information, including the National Best Bid and Offer (NBBO), as well as other important regulatory information such as short-sale restrictions and regulatory halts. In the highly fragmented world of US equities, the SIP is an easy way for people to get a view of the current state of the market.

FINRA TRF AND ODD LOT TRADES

FINRA TRF

Equity trades are executed on Public Exchanges (e.g., NASDAQ, BATS, NYSE, ARCA, etc.) and off the public exchanges in Dark Pools, Broker-Dealer internal crossing, and as Block Trades. Regulation National Market System (NMS) requires all trades to be reported. There



are currently three FINRA Trade Reporting Facilities (TRF) that are affiliated with registered national securities exchanges and provide FINRA members with a mechanism for the reporting of transactions affected otherwise than on an exchange.

Regulation NMS allows up to 10 seconds after the Trade execution time for the trade report to be sent to an exchange's TRF for publication. The delay can result in TRF Trade reports printed on the market data feed being out of the current NBBO.

Round Lot and Odd Lot

A round lot (or board lot) is a normal unit of trading of a security, which currently is 100 shares of stock in the US. Any quantity less than 100 shares is referred to as an odd lot. Odd lots are not subject to the Regulation NMS rules requiring execution to be within the current NBBO. Broker-dealers send odd lots to the exchange paying the most rebate per share and not the best execution price.

Note: Odd lot executions can create unrealistic high/low trade prices in an OHLC bar.

OFFICIAL OPEN/CLOSE IDENTIFICATION

The opening and closing trades are defined as the Opening and Closing Trade Cross on the Primary exchange. Identifying these trades is simple in theory; in actual practice, it has only become easy in the last few years as all the Primary Exchanges began publishing an Official Open and Close. algoseek identifies the Open/Close cross based on the Trade flags published with each trade.

Please refer to Appendix B for a detailed daily bars logic.

Opening Trade Identification

algoseek selects the Open Trade for an equity on the Primary Exchange from the priority order of trade types in this table.

Table 1: Open Trade from Trade Type

Priority	Open Order Type	Description
1	Primary Official Open	The Official Open trade on the Primary Exchange
2	Primary Opening Print	The Opening Print trades on the Primary Exchange
3	Regular Opening Trade	The largest regular size trade on the Primary Exchange from market Open to (Open + 15 minutes), which is normally 09:30 to 09:45 ET



4	First Trade Market Hours	The first regular trade on the Primary Exchange after the (Open + 15 minutes). This only applies to stocks with very small volumes
---	--------------------------	--

Closing Trade Identification

Since the exchanges have used different condition flags (or none at all) to identify the closing cross over the years, algoseek uses the following priority order to identify the Closing trade for a stock.

Table 2: Closing Trade for a Stock

Priority	Close Order Type	Description
1	Primary Official Close	The Official Close on Primary Exchange
2	Primary Closing Print	The sum of closing print trades on the Primary Exchange. The sum is needed as some exchanges (e.g. ARCA) used to print multiple closing prints instead of a single trade with the total size
3	Regular Closing Trade	The largest regular size trade on the Primary Exchange from Close to (Close + 5 minutes), which is normally 16:00 to 16:05:00 ET (except for half-days and early closes)
4	Last Trade before Close	The last trade on the Primary Exchange before the market Close

Daily BAR CALCULATIONS

Aggregated NBBO Bid/Ask Size

The Bid/Ask size in a bar field is the total aggregate from all the exchanges that have a matching price to the NBBO Bid/Ask price.

Quote Price Filter

When trading and quoting activities are inactive, for example, during extended trading hours or with an illiquid stock, bid prices can be extremely low, and ask prices can be extremely high. An exchange can also send a bad price, for example, a stock has a bid of \$12.05 then an exchange sends a bid of \$212.05.

To make Daily TAQ dataset usable for illiquid stocks and ETFs/ETNs, algoseek filters out extreme quotes by the following two criteria:



Bid price < (0.05 x average price of last 10 days)

Ask price > (10 x average price of last 10 days)

Separating Exchange and Finra Volume

The data has trade volume separated into fields:

ListedMarketHoursVolume: Trades done on the listed exchanges during market hours

ListedTotalVolume: Trades done on the listed exchanges during the whole trading day

FinraMarketHoursVolume: Trades done in Dark Pools, internally by Broker-Dealers, or on an Over-the-Counter (OTC) market reporting to FINRA during market hours

FinraTotalVolume: Trades done in Dark Pools, internally by Broker-Dealers, or on an Over-the-Counter (OTC) market reporting to FINRA during the whole trading day

The Volume data is separated to make it easy to understand the trading for either the public-listed exchanges or private non-public trading.

EXCHANGE MARKET HOURS

The US Stock Market trading hours are split into the pre-market, market, and post-market hours.

Pre-Market Hours: 04:00:00 to 09:30:00 (excluding)

Market Hours: 09:30:00 to 16:00:00 (excluding)

Post-Market Hours: 16:00:00 to 20:00:00

Note: Occasionally trade events are recorded several minutes after 20:00.

The “Open” normally refers to the Opening Cross at 09:30:00 ET. The “Close” normally refers to the Closing Cross at 16:00:00 ET.

Market Holidays and Early Closes

The stock market is closed for trading on most US holidays. For reference, algoseek publishes a list of historical holidays, which is available at `s3://us-equity-market-holidays/holidays.csv` (direct download link: <https://us-equity-market-holidays.s3.amazonaws.com/holidays.csv>).

Markets sometimes close early at 13:00:00 on the day before holidays such as Independence Day and Thanksgiving. You can download algoseek’s early close date and



time list from AWS S3 storage at

`s3://us-equity-market-holidays/earlycloses.csv` (or use a direct link us-equity-market-holidays.s3.amazonaws.com/earlycloses.csv).

DATA ORGANIZATION AND FILE FORMAT

algoseek provides Equity market data in plain text CSV files. The first row of CSV file is a fixed header and then rows of data corresponding to individual bars. By default, data is organized into one file per trading day. It means, for example, that all daily data for Mar 3, 2020, are stored in a separate CSV file

Table 3 (below) provides the name, base event, default value, brief description, and data type for each data field (column) in Equity Daily TAQ 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 3: CSV File Fields Schema

Field	Base Event	Type (Format)	Missing	Description
SecId	-	integer	Never	algoseek unique Security ID
TradeDate	-	string (yyyymmdd)	Never	Trading date in yyyymmdd format
Ticker	-	string	Never	Symbol name
Name	-	string	Blank	Name of equity security
PrimaryExchange	-	string	Blank	Primary listing exchange on this TradeDate
ISIN	-	string	Blank	ISIN as of this trade date. Optional.
OpenBidPrice	Quote	decimal	Never	NBBO Bid Price as of bar Open, (e.g., current price as of bar start)
OpenBidSize	Quote	integer	Never	Total Size from all exchanges with OpenBidPrice
OpenAskPrice	Quote	decimal	Never	NBBO Ask Price as of bar open (e.g. current price as of bar start)
OpenAskSize	Quote	integer	Never	Total Size from all Exchanges with NBBO OpenAskPrice
OpenTradePrice	Trade	decimal	Blank	Price of the first trade



OpenTradeSize	Trade	integer	Blank	Number of shares of the first trade
OpenTradeTime	Trade	string (timestamp)	Blank	Time of the first trade
HighBidPrice	Quote	decimal	Never	Highest NBBO bid price
HighBidSize	Quote	integer	Never	Total size from all exchanges with HighBidPrice
HighBidTime	Quote	string (timestamp)	Never	Time of highest NBBO bid price
HighAskPrice	Quote	decimal	Never	Highest NBBO ask price
HighAskSize	Quote	integer	Never	Total size from all exchanges with HighAskPrice
HighAskTime	Quote	string (timestamp)	Never	Time of highest NBBO ask price
HighTradePrice	Trade	decimal	Blank	Price of the highest trade
HighTradeTime	Trade	string (timestamp)	Blank	Time of the highest trade
LowBidPrice	Quote	decimal	Never	Lowest NBBO bid price of a bar
LowBidSize	Quote	integer	Never	Total Size from all exchanges with LowBidPrice
LowBidTime	Quote	string (timestamp)	Never	Time of the lowest bid
LowAskPrice	Quote	decimal	Never	Lowest NBBO Ask price of a bar
LowAskSize	Quote	integer	Never	Total size from all exchanges with LowAskPrice
LowAskTime	Quote	string (timestamp)	Never	Time of the lowest ask
LowTradePrice	Trade	decimal	Blank	Price of the lowest trade
LowTradeTime	Trade	string (timestamp)	Blank	Time of the lowest trade
CloseBidPrice	Quote	decimal	Never	NBBO Bid Price at bar Close
CloseBidSize	Quote	integer	Never	Total Size from all Exchange with CloseBidPrice



CloseAskPrice	Quote	decimal	Never	NBBO Ask Price at bar Close
CloseAskSize	Quote	decimal	Never	Total Size from all Exchange with CloseAskPrice
CloseTradePrice	Trade	decimal	Blank	Price of last Trade
CloseTradeSize	Trade	integer	Blank	Number of shares of last trade
CloseTradeTime	Trade	string (timestamp)	Blank	Time of the last Trade
MinSpread	Quote	decimal	Never	Minimum Bid-Ask spread size. This may be 0 if the market was crossed during the bar. If there is a negative spread due to back quote, make it zero
MaxSpread	Quote	decimal	Never	Maximum NBBO Bid-Ask spread in a bar
CancelSize	Trade	integer	Blank	Total shares canceled
NBBOQuoteCount	Quote	integer	0	Number of Bid and Ask NNBO quotes during the bar period
TradeAtBid	Quote Trade	integer	0	Sum of trade volume that occurred at or below the bid (a trade reported/ printed late can be below the current bid)
TradeAtBidMid	Quote Trade	integer	0	Sum of trade volume that occurred between the bid and the midpoint: $\text{TradeAtBidMid} = (\text{Trade Price} > \text{NBBO Bid}) \& (\text{Trade Price} < \text{NBBO Mid})$
TradeAtMid	Quote Trade	integer	0	Sum of trade volume that occurred at mid. $\text{TradePrice} = \text{NBBO MidPoint}$
TradeAtMidAsk	Quote Trade	integer	0	Sum of ask volume that occurred between the mid and ask. $\text{TradeAtMidAsk} = (\text{Trade Price} > \text{NBBO Mid}) \& (\text{Trade Price} < \text{NBBO Ask})$
TradeAtAsk	Quote Trade	integer	0	Sum of trade volume that occurred at or above the Ask
TradeAtCrossOrLocked	Quote Trade	integer	0	Sum of trade volume for the bar when NBBO is locked or crossed. Locked is Bid = Ask Crossed is Bid > Ask



ListedMarketHoursVolume	Trade	integer	0	Public Listed exchanges trading volume during regular market hours only
ListedMarketHoursTrades	Trade	integer	0	Number of trades during regular market hours in public listed exchanges
ListedTotalVolume	Trade	integer	0	Public Listed exchanges trading volume for the whole day (includes pre, regular, and post-market)
ListedTotalTrades	Trade	integer	0	Public Total number of trades for the trade date (includes pre-, regular, and post-market) in listed exchanges
FinraMarketHoursVolume	Trade	integer	0	FINRA/TRF trading volume during regular market hours only (normal trade day is 09:30:00 to 16:00:00 ET). FINRA/TRF represents off-exchange trading.
FinraMarketHoursTrades	Trade	integer	0	The number of FINRA/TRF trades during regular market hours. FINRA/TRF represents off-exchange trading.
FinraTotalVolume	Trade	integer	0	FINRA/TRF trading volume for the whole day (includes pre-, regular, and post-market)
FinraTotalTrades	Trade	integer	0	Total number of FINRA/TRF trades for the trade date (includes pre, regular, and post-market). FINRA/TRF represents off-exchange trading.
UptickVolume	Trade	Integer	0	Total number of shares traded with upticks during the bar. Uptick = (Trade Price > Last Trade Price)
DowntickVolume	Trade	Integer	0	Total number of shares traded with downticks during the bar. Downtick = (Trade Price < Last Trade Price)
RepeatUptickVolume	Trade	Integer	0	Total number of shares where trade price is the same (repeated) and last price change was up during the bar. Repeat Uptick = (Trade Price == Last Trade Price) & (Last Tick Direction == Up)



RepeatDowntickVolume	Trade	Integer	0	Total number of shares where trade price is the same (repeated) and last price change was down during the bar. Repeat Downtick = (Trade Price == Last Trade Price) & (Last Tick Direction == Down)
UnknownTickVolume	Trade	Integer	0	When the first trade of the day takes place, the tick direction is “unknown” as there is no previous trade to compare it to. This field is the volume of the first trade after 4 AM and acts as an initiation value for the tick volume directions.
MarketVWAP	Trade	decimal	Blank	Volume weighted average price during regular market hours, normally between 09:30:00 and 16:00:00 ET plus the Opening and Closing Cross (which may be after 16:00:00 ET).
DailyVWAP	Trade	decimal	Blank	Volume weighted average price for the whole day including pre, regular, and post-market trades.
TradeToMidVolWeight	Quote Trade	decimal	Blank	Indicator for the bar period showing the sum difference between each trade’s price and NBBO midpoint at the time of the trade weighted by volume. It returns a positive or negative number indicating buying or selling pressure. Note: Blank if no Trades. FINRA reported trades are not included
TradeToMidVolWeightRelative	Quote Trade	decimal	Blank	Indicator for the bar period showing the sum difference between each trade’s price and NBBO midpoint at the time of the trade relative to the spread and weighted by volume. It returns a positive or negative number indicating buying or selling pressure. Note: Blank if no trades. FINRA reported trades are not included.
TimeWeightBid	Quote	decimal	Blank	Time-weighted average price of National Best Bid during the bar period
TimeWeightAsk	Quote	decimal	Blank	Time-weighted average price of National Best Ask during the bar period



Bar Notes

MarketVWAP and DailyVWAP: Market volume weighted average price (VWAP) is calculated as

$$\text{sum(TradePrice} \times \text{TradeSize)} / \text{sum(TradeSize)}$$

where

$$\text{sum(TradeSize)} = \text{ListedMarketHoursVolume} + \text{FinraMarketHoursVolume}$$

TradeToMidVolWeight, TradeToMidVolWeightRelative: volume-weighted trade to the midpoint is calculated as the following sum over all trades during the bar

$$\frac{\text{sum(Trade_Shares} \times (\text{Trade_Price} - \text{NBBOMidpoint}))}{\text{sum(Trade_Shares)}}$$

Similarly, volume-weighted relative trade to the midpoint

$$\frac{\text{sum(Trade_Shares} \times (\text{Trade_Price} - \text{NBBOMidpoint}) / \max(1, \text{NBBOSpread}))}{\text{sum(Trade_Shares)}}$$

where midpoint and spread values are calculated based on the last NBBO

$$\text{NBBOMidpoint} = (\text{NBBOBid_InPennies} + \text{NBBOAsk_InPennies}) / 2$$

$$\text{NBBOSpread} = \text{NBBOAsk_InPennies} - \text{NBBOBid_InPennies}$$

If Bid == Ask, then it is assumed the midpoint of the Bid/Ask is that price. If the market is crossed (NBBO Bid > NBBO Ask), then it is not possible to know what the correct price is so the last good NBBO Bid and Ask (including the Bid == Ask case) will be used.

TimeWeightBid, TimeWeightAsk: time-weighted bid and ask are calculated with

$$\text{sum(Price}_{\{n\}} \times (\text{Price}_{\{n+1\}_Time} - \text{Price}_{\{n\}_Time}) / \text{Bar_Duration}$$

where Price₀ is the bar open price.



APPENDIX A. FREQUENTLY ASKED QUESTIONS

Why are Trade Prices often inside the Bid Price to Ask Price range?

The Low/High Bid/Ask is the low and high NBBO price for the bar range. Very often, a trade may not occur at these prices, as the price may only last a few seconds or executions are being crossed at mid-point due to hidden order types that execute at mid-point or as price improvement over current Bid/Ask.

Which fields give the total volume for the regular market hours?

Total volume can be calculated as

$$\text{TotalVolume} = \text{ListedMarketHoursVolume} + \text{FinraMarketHoursVolume}$$

When the ticker doesn't have any trade, the columns OpenTradePrice and OpenTradeSize are null. Where does the value for the CloseTradePrice and CloseTradeSize come from?

Null values at columns OpenTradePrice (CloseTradePrice) and OpenTradeSize (CloseTradeSize) for low liquid symbols are caused by the absence of trades around the Open (Close) time for such tickers. At the same time, trades can appear later (earlier) in the SIP feed, that is why we only see CloseTradePrice (OpenTradePrice) and CloseTradeSize (OpenTradeSize).

Why are open/close prices in algoseek's datasets different from other sources?

algoseek provides official opening/closing cross price based on stock's primary exchange, event flags, time, and volume.

Most Data vendors simply use the first/last trade, or the opening/closing cross from a random exchange as opening/closing price.

Why are time-based columns not properly recognized when I try importing data to Excel?

Older versions of Excel will automatically convert the TimeBarStart field into an Excel format timestamp, but this fails when TimeBarStart is HHMMSSmmm (millisecond) or HHMMSSmmmmiiinnn (nanosecond). For timestamp with the nanosecond (millisecond) format, import the data using Excel "From Text" option and set the data type for column "TimeBarStart" to "Text", so Excel does not automatically try to convert it.



APPENDIX B. BAR CALCULATIONS FROM TRADE AND QUOTE EVENTS

This section describes the logic for minute bar calculations based on events from the Trade and Quote dataset. Please also refer to the Equity Trade and Quote Guide for more details of data fields and condition flags used.

In this section, Market hours refer to 9:30:00 (including) - 16:00:00 (excluding) for regular trading days and to 9:30:00 (including) - 13:00:00 (excluding) when there was an early close.

Opening Trade

The Opening Trade is selected from the following priority order list:

- the last event with tOfficialOpen flag from the Primary Exchange
- the last event with tOpeningPrint flag from the Primary Exchange
- a Regular Open from the Primary Exchange
- a Regular First from the Primary Exchange

Regular Open

The Regular Open corresponds to the first event satisfying all the conditions below:

- not an Open/Close event (tOpeningPrints, tClosingPrints, tOfficialClose and tOfficialOpen flags)
- not having tExtendedHours flag set
- the largest volume between 09:30:00 (including) and 09:40:00 (including)

Regular First

Regular First is the first event satisfying all the conditions below:

- not an Open/Close event (tOpeningPrints, tClosingPrints, tOfficialClose and tOfficialOpen flags)
- not having either of the following flags set: tDerivativelyPrices, tStockOption, tAveragePrice, tRule155
- not having tOddLot flag set
- not having tExtendedHours flag set
- the event time is within market hours



Closing Trade

The Closing Trade is selected from the following priority order list:

- the first event with tOfficialClose flag from the Primary Exchange
- the first event with tClosingPrint flag from the Primary Exchange
- a Regular Close from the Primary Exchange
- a Regular Last from the Primary Exchange

Regular Close

The Regular Close is the last event satisfying all the condition below:

- not an Open/Close event (tOpeningPrints, tClosingPrints, tOfficialClose and tOfficialOpen flags)
- not having tExtendedHours flag set
- the largest volume between 16:00:00 (including) and 16:05:00 (including) for regular trading days or between 13:00:00 (including) and 13:05:00 (including) for an Early Close

Regular Last

Regular Last is the last event satisfying all the conditions below:

- not an Open/Close event (tOpeningPrints, tClosingPrints, tOfficialClose and tOfficialOpen flags)
- not having either of the following flags set: tDerivativelyPrices, tStockOption, tAveragePrice, tRule155
- not having tOddLot flag set
- not having tExtendedHours flag set
- the event time is within market hours

Daily High

Daily High is the first event satisfying all the conditions below:

- not an Open/Close event (tOpeningPrints, tClosingPrints, tOfficialClose and tOfficialOpen flags)
- not having either of the following flags set: tDerivativelyPrices, tStockOption, tAveragePrice, tRule155
- not having tOddLot flag set
- not having tExtendedHours flag set
- the exchange is not FINRA
- the event with the largest price within market hours



Daily Low

Daily Low corresponds to the first event satisfying all the conditions below:

- not an Open/Close event (tOpeningPrints, tClosingPrints, tOfficialClose and tOfficialOpen flags)
- not having either of the following flags set: tDerivativelyPrices, tStockOption, tAveragePrice, tRule155
- not having tOddLot flag set
- not having tExtendedHours flag set
- the exchange is not FINRA
- the event with the smallest price within market hours

Quote Data

Excluded data

- Exclude any quote with Price = 0
- Exclude any quote with Quantity = 0

You should also exclude any event with one or more flags listed in Table 4.

Table 4: Flags for Quote Events to be Excluded During Bar Calculations

Quote Events	
Bit Mask Position	Flags
3	qClosing
4	qNewsDissemination
5	qNewsPending
6	qTradingRangeIndication
7	qOrderImbalance
13	qResume

Included data

- Include only NBBO quotes



You should only include events with one or more flags listed in Table 5. If the event has any of the exclude flags enabled, it is not included. If the event does not contain any flags from the include list, it is not included in bar calculations.

Table 5: Flags for Quote Events to be Included During Bar Calculations

Quote Events	
Bit Mask Position	Flags
0	qRegular
1	qSlow
2	qGap
11	qOpeningQuote
21	qFastTrading

Price validation

Additionally, you should filter out test quote events using the following approach:

- 1) For symbols with price history (last 10 trading days):

$$\text{MinPrice} = 0.05 * \text{AveragePrice}$$

$$\text{MaxPrice} = 10 * \text{AveragePrice}$$

- 2) For new symbols (no price history):

$$\text{MinPrice} = 0.03$$

$$\text{MaxPrice} = 19998$$

If a Quote price is lower than MinPrice or higher than MaxPrice - the event is excluded.

Note: We do not recommend applying price filtering for Trade events.