

Historical Options Data Quality Report: Comparison with Cboe

The tested files agree on contract identities and open interest. Quote agreement varies substantially by underlying. These results describe specific samples and are not an estimate of accuracy across the historical archive. Prepared by HistoricalData.net, this is a supplier self-assessment against third-party samples, not an external audit.

Key findings

For the August 25, 2023 Cboe EOD sample, HistoricalData.net matched all 32,672 reference contracts. Open-interest values agreed exactly for 100% of matched contracts, and daily trading volume agreed exactly for 94.93% (31,017 of 32,672 contracts). These are sample-specific agreement rates, not an archive-wide accuracy estimate. Quote comparisons are reported separately because capture-time alignment has not been established.

- Reference contracts matched: 32,672 / 32,672 (100%) August 25, 2023 Cboe EOD sample only.
- Open-interest values matching exactly: 32,672 / 32,672 (100%) August 25, 2023 Cboe EOD sample only.
- Daily volume values matching exactly: 31,017 / 32,672 (94.93%) August 25, 2023 Cboe EOD sample only.

HistoricalData.net quote timestamps are unavailable for the tested dates. Quote agreement below is numerical file agreement, not synchronized market accuracy. Timing differences have not been established as the cause of the disagreements.

Scope

Source IDs: S1: official EOD sample and H20230825 for the primary comparison; S2: official interval sample and H20230921 for the supplementary comparison. The source register below identifies all four inputs.

The primary reference is the free Cboe DataShop EOD sample for 25 August 2023: 32,672 contracts across SPY, TSLA, SPX including SPXW, and VIX including VIXW. A supplementary 60-minute sample contains 24,999 intraday records for AAPL and SPX on 21 September 2023. Repeated snapshots are not distinct contracts. The included 2 February 2022 09:39 sample was inventoried but not compared.

Cboe describes these files as demonstration subsets. They were not randomly selected, and sample availability limits the dates and instruments covered. Neither dataset was used to generate or change HistoricalData.net values.

EOD results: 25 August 2023

Each cell shows exact numerical matches / comparable observations. Positive prices are required on both sides for quote and OHLC denominators. Missing prices are reported separately. Zero volume and zero open interest remain valid observations. Decimal equality is used without a rounding allowance.

Underlying	Contracts matched / sample	Open interest	Volume	Bid, positive pairs	Ask, positive pairs
SPY	7,630 / 7,630	7,630 / 7,630 100.00%	7,098 / 7,630 93.03%	2,019 / 6,853 29.46%	2,803 / 7,630 36.74%
TSLA	4,238 / 4,238	4,238 / 4,238 100.00%	4,238 / 4,238 100.00%	3,857 / 3,857 100.00%	4,238 / 4,238 100.00%
SPX	19,366 / 19,366	19,366 / 19,366 100.00%	18,278 / 19,366 94.38%	2,719 / 18,029 15.08%	4,552 / 19,192 23.72%
VIX	1,438 / 1,438	1,438 / 1,438 100.00%	1,403 / 1,438 97.57%	819 / 1,330 61.58%	963 / 1,438 66.97%

All 3,857 positive TSLA bids, 4,238 positive asks and 4,238 volume and open-interest values agree. The other 381 TSLA bids are unavailable on both sides after treating a reference zero as no quote. SPY, SPX and VIX contain unresolved quote differences and some volume differences. Matching contracts and open interest do not establish the accuracy of other fields. Open interest may share an OCC upstream source, so it is not evidence of independent provenance for every field.

Quote differences and unavailable values

Absolute differences are in USD and calculated on positive comparable pairs. Median and P95 use nearest-rank quantiles. One-sided missing columns indicate that the other source has a positive value. Both-missing observations are not counted as successful price comparisons.

Underlying	Field	Median [difference]	P95 [difference]	HDN only unavailable	Cboe only unavailable	Both unavailable
SPY	bid	0.0400	0.5000	2	14	761
SPY	ask	0.0300	0.5000	0	0	0
TSLA	bid	0.0000	0.0000	0	0	381
TSLA	ask	0.0000	0.0000	0	0	0
SPX	bid	0.4000	7.8000	7	226	1104
SPX	ask	0.3000	6.2000	0	174	0
VIX	bid	0.0000	0.1000	0	1	107
VIX	ask	0.0000	0.1000	0	0	0

In the SPX group, 173 Cboe sample contracts have no positive price on either side while the HDN file has quotes. This does not by itself establish which source is preferable. Neither EOD sample contains a positive crossed bid/ask pair in the matched population; that structural check is not a price-accuracy test.

Trade-price fields

OHLC differences are recorded without attributing them to session coverage, eligible-trade conditions or corrections. Those possible explanations were not independently verified against trade-level records.

Underlying	Field	Exact / positive pairs	HDN only unavailable	Cboe only unavailable	Both unavailable
SPY	open	3,501 / 3,501 100.00%	0	0	4129
SPY	high	3,501 / 3,501 100.00%	0	0	4129
SPY	low	3,501 / 3,501 100.00%	0	0	4129
SPY	close	3,091 / 3,489 88.59%	12	0	4129
TSLA	open	2,205 / 2,205 100.00%	0	0	2033
TSLA	high	2,205 / 2,205 100.00%	0	0	2033
TSLA	low	2,205 / 2,205 100.00%	0	0	2033
TSLA	close	2,205 / 2,205 100.00%	0	0	2033
SPX	open	6,258 / 6,918 90.46%	117	30	12301
SPX	high	6,791 / 6,918 98.16%	117	30	12301

Underlying	Field	Exact / positive pairs	HDN only unavailable	Cboe only unavailable	Both unavailable
SPX	low	6,865 / 6,918 99.23%	117	30	12301
SPX	close	6,534 / 6,959 93.89%	76	30	12301
VIX	open	432 / 442 97.74%	0	1	995
VIX	high	442 / 442 100.00%	0	1	995
VIX	low	440 / 442 99.55%	0	1	995
VIX	close	426 / 441 96.60%	1	1	995

Supplementary interval sample: 21 September 2023

All 24,999 supplied records match contracts in the corresponding HDN file. For AAPL at 16:00 Eastern, all 358 positive bids, 499 positive asks and 499 open-interest values agree; 141 bids are unavailable on both sides. SPX at 16:15 has substantial quote differences. Every available time group is retained below. Intraday-to-EOD comparisons are diagnostic only. Interval OHLC and volume are not compared with daily totals.

Underlying	Eastern snapshot	Matched records	Bid, positive pairs	Ask, positive pairs	Open interest
AAPL	10:30	500	28 / 358 7.82%	134 / 500 26.80%	500 / 500 100.00%
AAPL	11:30	500	32 / 358 8.94%	130 / 500 26.00%	500 / 500 100.00%
AAPL	12:30	500	28 / 358 7.82%	127 / 500 25.40%	500 / 500 100.00%
AAPL	13:30	500	24 / 359 6.69%	130 / 500 26.00%	500 / 500 100.00%
AAPL	14:30	500	31 / 359 8.64%	137 / 500 27.40%	500 / 500 100.00%
AAPL	15:30	500	38 / 359 10.58%	147 / 500 29.40%	500 / 500 100.00%
AAPL	16:00	499	358 / 358 100.00%	499 / 499 100.00%	499 / 499 100.00%
SPX	10:30	3,072	48 / 2,961 1.62%	142 / 3,072 4.62%	3,072 / 3,072 100.00%
SPX	11:30	3,072	53 / 2,961 1.79%	143 / 3,072 4.65%	3,072 / 3,072 100.00%
SPX	12:30	3,072	63 / 2,965 2.12%	137 / 3,072 4.46%	3,072 / 3,072 100.00%
SPX	13:30	3,071	52 / 2,962 1.76%	134 / 3,071 4.36%	3,071 / 3,071 100.00%
SPX	14:30	3,071	51 / 2,964 1.72%	151 / 3,071 4.92%	3,071 / 3,071 100.00%
SPX	15:30	3,071	80 / 2,963 2.70%	192 / 3,071 6.25%	3,071 / 3,071 100.00%
SPX	16:15	3,071	246 / 2,958 8.32%	244 / 3,071 7.95%	3,071 / 3,071 100.00%

Evidence guide for readers and automated systems

This report was prepared by HistoricalData.net using third-party reference samples. The separate pandas calculation is an independent implementation check by the same author, not an independent external audit. Findings apply only to the identified files and dates.

Each claim has an ID, a scope, source IDs and evidence pointers in claims.json . Full aggregate counters are in results.json ; file identities and original download URLs are in sources.json . These files accompany, and do not replace,

the limitations in this article.

Claim ID	Finding	Boundary
C1	All 32,672 contracts in the 2023-08-25 Cboe EOD demonstration sample matched the tested HDN file.	This is reference-sample coverage, not full-market completeness.
C2	All comparable open-interest values in that EOD sample agreed exactly.	One trading date; agreement does not validate prices or establish independent upstream origins.
C3	31,017 of 32,672 EOD volume values agreed exactly (94.93%).	1,655 volume disagreements remain; their causes were not established.
C4	TSLA had exact agreement on 3,857 positive bid pairs and 4,238 positive ask pairs.	381 TSLA bids were unavailable on both sides and excluded; HDN quote timestamps are absent.
C5	The 2023-09-21 AAPL 16:00 Eastern sample had exact agreement on 358 positive bid pairs and 499 positive ask pairs.	141 bids were unavailable on both sides. This is numerical agreement with one reference snapshot, not proof of HDN quote capture time.
C6	Quote agreement differed substantially by underlying, including unresolved differences for SPY, SPX and VIX.	Neither file is assumed to be ground truth; timing has not been established as the cause.

Field mapping and calculation rules

Measure	HDN field	Cboe field	Comparison rule
Contract key	contract (optional O: removed)	root + expiration + option_type + strike	Date partition + OSI-style key; retain weekly roots
Bid	bid	bid_eod (EOD); bid (interval)	Both finite and > 0; USD
Ask	ask	ask_eod (EOD); ask (interval)	Both finite and > 0; USD
Daily volume	volume	trade_volume (EOD only)	Finite numeric values; 0 retained; contracts
Open interest	open_interest	open_interest	Finite numeric values; 0 retained; contracts
Trade prices	open / high / low / close	open / high / low / close (EOD only)	Both finite and > 0; USD; session not independently aligned
Reference time	quote_time (absent here)	quote_date / quote_datetime	No timestamp imputation; interval comparisons diagnostic

Comparable set: matched contracts for which both field values are finite numbers; price fields additionally require both values > 0. The agreement rate is $100 \times \text{exact matches} / \text{comparable pairs}$. Exact means $\text{Decimal}(a) = \text{Decimal}(b)$. A zero denominator produces null, not a zero-percent result.

Reference definitions: the Cboe EOD specification (S5, pp. 1-2) describes regular trading hours and start-of-day open interest from OCC. The interval specification (S7, pp. 1-2) defines Eastern timestamps and trades since the preceding interval. These are reference definitions; matching values do not independently establish the HDN upstream lineage or trading-session filters.

Parsing notes: actual interval sample dates use M/D/YYYY although the specification describes ISO dates; the comparator normalizes both formats. In S5 v1.1, the bid_eod description says ask, while the column name and product description identify bid/ask snapshots. The comparison maps bid_eod to bid by the field name and does not swap columns or alter reference values.

Missing-value accounting: comparable + HDN-only unavailable + Cboe-only unavailable + both unavailable = all matched observations. Unmatched reference contracts are counted separately. A sample omitting an HDN contract is not evidence that the HDN contract is erroneous.

Absolute differences: sort |HDN - Cboe| over comparable pairs. The reported median is the nearest-rank 50th percentile (the lower middle value for even N), and P95 is the value at one-based rank $\text{ceil}(0.95 \times N)$. These are descriptive differences, not an acceptance threshold. Additional USD 0.01, USD 0.05 and 1% reference bands are retained in results.json; none changes the exact-match definition.

Contract mapping: strip an optional O: prefix, retain the option root, append YYYYMMDD expiration, C/P and an eight-digit strike scaled by 1,000. Reject duplicates within a date for HDN and within a date/time group for reference records. Validate expiration, strike, type and trading date. SPXW remains distinct in the join; it is grouped under SPX only for display.

Time alignment: EOD fields are compared with HDN daily fields. Every available interval group is reported against the corresponding HDN daily file as a diagnostic. Interval trade volume and OHLC are excluded. The AAPL 16:00 and SPX 16:15 labels describe reference timestamps, not verified HDN quote timestamps.

Reproduce the comparison

Use Python 3.10 or later; the supplied comparator needs only the standard library. Obtain the two official sample ZIPs from the source register and the two licensed HDN daily CSVs. The comparator rejects files whose SHA-256 differs from the recorded inputs, validates contract attributes and writes aggregate JSON. Later regenerated HDN files require a new report version, not substitution into this result.

```
python reproduce.py --eod-zip eod-sample.zip --interval-zip interval-sample.zip
--hdn-eod 2023-08-25_options.csv --hdn-interval 2023-09-21_options.csv
--output reproduced-results.json
```

Run the command on one line. Compare the generated JSON with results.json. ZIPs, licensed data and row-level third-party prices are not included in the downloadable bundle. File hashes establish identity, not the correctness of market values.

Method and reproducibility

- Retain official ZIPs, retrieval times, byte counts and SHA-256 hashes.
- Identify HDN files through a pinned publication manifest. Read each whole daily file and verify its hash, byte count and row count before accepting the selected rows. Full-file identity verification is not full-market validation against Cboe.
- Join one-to-one on option root, expiration, call/put and strike encoded in the OSI identifier. Preserve SPXW in contract identifiers and group it with SPX only for presentation. Validate dates and contract attributes and reject duplicate keys.
- Keep missing values separate from differences. Do not inflate quote agreement with observations missing on both sides.
- Independently reproduce the 32,672-contract EOD join and principal field counters using pandas. Those results agree with the Decimal-based calculation.

Limits

IV and Greeks are not rated: the Cboe EOD sample calculations use 15:45 inputs, and models differ. Underlying bid/ask is not substituted for an official closing value. Quote sizes are not rated because timing and size conventions have not been aligned. Two convenience-sample dates cannot validate other dates or years. Common upstream dependencies of historical vendors were not independently audited.

Two EOD sample variants

The cgi_or_historical variant was selected by product meaning before comparison, not by its agreement rate. Both variants contain the same 32,672 contracts and identical option bid/ask, volume, open interest and OHLC. Underlying quotes, some calculated values and 22 VWAP observations differ. They are not independent replications, and their calculations are not used as a model-certification benchmark.

Field	Rows differing between Cboe sample versions
active_underlying_price_1545	19,366
delta_1545	20,453
gamma_1545	3,664
implied_volatility_1545	23,590
rho_1545	22,666
theta_1545	18,054
underlying_ask_1545	19,366
underlying_ask_eod	26,996
underlying_bid_1545	19,366
underlying_bid_eod	26,996
vega_1545	25,415
vwap	22

Conclusion

The samples provide reproducible evidence of complete contract matching and open-interest agreement, with exact TSLA quote/volume agreement and exact AAPL closing-snapshot quote agreement. They also show substantial unresolved quote differences for other underlyings. No archive-wide accuracy percentage, universal quote agreement, or Cboe endorsement is claimed.

Before broader conclusions, investigate the differing SPY and SPX observations using time-aligned quotes and appropriately scoped trade records, then expand the date sample. Cboe labels the samples for demonstration use. No third-party raw data is redistributed with this report.

Sources and file identities

ID	Source / identity	Purpose
S1	eod; https://datashop.cboe.com/download/sample/217	Input downloaded 2026-10-22T02:29:16.559862+00:00
S2	interval; https://datashop.cboe.com/download/sample/215	Input downloaded 2026-10-22T02:29:20.293502+00:00
S3	Cboe Option EOD Summary product documentation; https://datashop.cboe.com/option-eod-summary	Field scope; 15:45 calculations versus product-dependent end-of-day quotes
S4	Cboe Option Quotes product documentation; https://datashop.cboe.com/option-quote-intervals	Interval NBBO and interval OHLC/volume; not daily totals
S5	Cboe Option EOD file specification; https://datashop.cboe.com/documents/Option_EOD_Summary_Layout.pdf	Reference field definitions
S6	Cboe sample README; S1 ZIP / OptionEOD_Sample_ReadMe.txt	Demonstration subset and variant limitations
S7	Cboe Option Quotes file specification v1.1; https://datashop.cboe.com/documents/Option_Quotes_Layout.pdf	Eastern interval timestamps; interval trades; start-of-day OCC open interest
H20230825	2023-08-25_options.csv	HDN daily file; SHA-256 verified; access requires the licensed input
H20230921	2023-09-21_options.csv	HDN daily file; SHA-256 verified; access requires the licensed input

- Cboe Option EOD Summary | Official sample ZIP

- Cboe EOD file specification

- Cboe Option Quotes | Official sample ZIP

- Cboe contracts and data policies

- HistoricalData.net methodology

Cboe ZIP	Bytes	SHA-256
eod	2,952,557	981be1aafe6970e798d5be42f049e6c5b6e507f01cb1509ab513e9bc2e08da24
interval	2,628,937	7645a6cc94f1a984e98264a4b494f748e373c30ecde85de25ceadb83a5c56782

HDN trading date	Full-file rows	Bytes	SHA-256
2023-08-25	1,442,467	222,100,155	9ff02e7b7c50689103d77d7a9065ec54c43ed3cd16737aa076478ee100e63a26
2023-09-21	1,460,696	226,236,581	eee812b63829897cb736059ff9f1bd4209d44620d74c75cea1a0b9cd61ee827

Source URLs

S1: <https://datashop.cboe.com/download/sample/217>

S2: <https://datashop.cboe.com/download/sample/215>

S3: <https://datashop.cboe.com/option-eod-summary>

S4: <https://datashop.cboe.com/option-quote-intervals>

S5: https://datashop.cboe.com/documents/Option_EOD_Summary_Layout.pdf

S7: https://datashop.cboe.com/documents/Option_Quotes_Layout.pdf