US Equity Security Master and Lookup Files Guide

version 1.7 (Oct 2023)

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

The Equity Security Master File is a single data file containing all the listed and delisted equity securities including stocks, ETFs, ETNs, ADRs, stock warrants, preferred stocks, etc., from 2007 to present with summary information and industry identifiers.

The Security Master File is organized based on algoseek's unique identifier called Security ID (SecId), which remains unchanged during name or ticker changes.

Lookup Files are available for converting Ticker and FIGI identifiers to SecId.

DATA ORGANIZATION AND FILE FORMAT

List of Data Fields

Security Master File is provided as a single file (**equity_security_master.csv**) in CSV format, with each row corresponding to an individual security record.

Security Master File is updated on a daily basis and can be used to backtrack any historical modifications to a SecId or to restore point-in-time data.

Table 1 demonstrates the full list of data fields in the Security Master File with sample contents for a few SecIds, in which rows and columns are inverted for the convenience of the document display. For instance, one can see that SecId 33008 has changed its ticker and company name three times during 2007-2020.

Table 1: Sample Data from Security Master File

SecId	33008	33449	549535
Tickers	AA; ARNC; HWM	AAPL	SPY
TickersStartToEndDate	20070103:20161031; 20161101:20200331; 20200401:29991231	20070103:29991231	20070103:29991231
Name	Alcoa Inc.; Arconic Inc; Howmet Aerospace Inc	Apple Computer Inc.; Apple Inc	SSGA SPDR S&P 500
NameStartToEndDate	20070103:20160630; 20160701:20200331; 20200401:29991231	20070103:20070110; 20070111:29991231	20070103:29991231
ISIN	US0138171014;	US0378331005	US78462F1030

	US0138175072; US03965L1008; US4432011082		
ISINStartToEndDate	20070103:20161005; 20161006:20161031; 20161101:20200331; 20200401:29991231	20070103:29991231	20070103:29991231
ListStatus	L	L	L
SecurityDescription	Equity Shares	Equity Shares	Exchange Traded Fund
USIdentifier	013817101; 013817507; 03965L100; 443201108	037833100	78462F103
USIdentifierStartToEnd Date	20070103:20161005; 20161006:20161031; 20161101:20200331; 20200401:29991231	20070103:29991231	20070103:29991231
PrimaryExchange	NYSE	NASDAQ	AMEX; ARCA
PrimaryExchangeStartToEndDate	20070103:29991231	20070103:29991231	20070103:20081128; 20090224:29991231
SEDOL	BD3D9G5	2046251	2840215
Sic	3350	3571	
Sector	Manufacturing	Manufacturing	
Industry	Rolling Drawing & Extruding Of Nonferrous Metals	Electronic Computers	
FIGI	BBG000B9WH86; BBG000B9WH86; BBG00DYNJGH9; BBG00DYNJGH9	BBG000B9XRY4	BBG000BDTBL9

Table 2 below summarizes the name, brief description, and data type for each data field (column) in Equity Security Master File. The table column "Missing" indicates a default value or behavior in case the data field value is not present or cannot be determined where "Never" means that a value is always present in the data field.

Table 2: CSV File Fields Schema for Security Master

Field	Type (Format)	Missing	Description
SecId	integer	Never	algoseek unique security identifier
Tickers	string (tricker1; ticker2;…)	Never	List of symbol names used. If security had its ticker changed, the field will have multiple tickers separated by a semicolon ";"
TickersStartToEndDate	string (yyyymmdd: yyyymmdd;…)	Never	Start and end dates for each ticker. EndDate = 20991231 when the ticker is still being used
Name	string (name1; name2;…)	Never	List of security names used. If security had its name changed, the field will have multiple tickers separated by a semicolon ";"
NameStartToEndDate	string (yyyymmdd: yyyymmdd;…)	Never	Start and end dates for each name. EndDate = 20991231 when the name is still being used
ISIN	string (ISIN1; ISIN2;…)	Blank	List of ISIN codes used. If security had its ISIN changed, the field will have multiple tickers separated by a semicolon ";"
ISINStartToEndDate	string (yyyymmdd: yyyymmdd;…)	Blank	Start and end dates for each ISIN. EndDate = 20991231 when the ISIN is still being used
ListStatus	string	Blank	Current list status: A = Announced, D = Delisted, L = Listed
SecurityDescription	string	Blank	Current Security Description
USIdentifier	string (name1; name2;…)	Blank	List of USIdentifiers for US securities used. If security had its USIdentifier changed, the field will have multiple identifiers separated by a semicolon ";"
USIdentifierStartToEndDate	string (yyyymmdd: yyyymmdd;…)	Blank	Start and end dates for each USIdentifier. EndDate = 20991231 when the USIdentifier is still being used
PrimaryExchange	string (exchange1; exchange2;…)	Never	List of Primary Exchange(s). If security had its Primary Exchange changed, the field will have multiple exchange names separated by a semicolon ";"
PrimaryExchange StartToEndDate	string (yyyymmdd: yyyymmdd;…)	Never	Start and end dates for each Primary Exchange. EndDate = 20991231 when the Primary Exchange is still being used

SEDOL	string	Blank	Current Stock Exchange Daily Official List
Sic	integer	Blank	Current Standard Industrial Classification code
Sector	string	Blank	Current SIC Sector
Industry	string	Blank	Current SIC Industry
FIGI	string (FIGI1; FIGI2;…)	Blank	Financial Instrument Global Identifier. If security had its FIGI changed, the field will have multiple exchange names separated by a semicolon ";"

Note: The SEDOL (Stock Exchange Daily Official List) column includes only historical data. It is currently deprecated and no longer updated.

The same security ID can have multiple tickers, ISINs, Primary Exchanges, etc., during different periods of time in history. In algoseek's Equity Security Master File, all the historical values of a field are listed in text using a semicolon to separate the values.

For the StartToEndDate field, a format of yyyymmdd:yyyymmdd is used to indicate the start and end date of a period of time, and a semicolon is used to separate the different periods. For example:

```
20111108:20141210;20171108:20991231
```

in which the end date value 20991231 implies that the range is ongoing and no end date has been set.

Security Type

The "SecurityDescription" column in the Security Master File provides the type of security. Table 3 below includes a list of available security types and their brief descriptions.

Table 3: Security Types

SecurityDescription	Details
Equity Shares	Common stock
Structured Product	A pre-packaged investment that normally includes assets linked to interest plus one or more derivatives
Bond	A fixed-income instrument representing a loan made by an investor to a borrower
Exchange Traded Fund	A type of security that involves a collection of securities that often tracks an underlying index
Depository Receipts	Negotiable certificates issued by a bank representing shares in a foreign company traded on a local stock exchange

Preference Share	Shares of a company's stock with dividends that are paid out to shareholders before common stock dividends are issued
Units	Securities representing a collection of different (and usually related) shares
Warrants	Securities that give the holder the right to purchase a company's stock at a specific price and at a specific date
Preferred Security	See Preference Share
Tradeable Rights	A security representing an invitation to existing shareholders to purchase additional new shares in the company
Stapled Security	A type of financial instrument consisting of two or more securities that are contractually bound to form a single salable unit; they cannot be bought or sold separately.
Contingent Value Rights	A type of security ensuring that the shareholders get certain benefits if a specific event occurs, usually within a specified time frame
When Issued	A transaction that is made conditionally because a security has been authorized but not yet issued

Industry Sector

Securities are categorized into industry sectors. The Sector field can be one of the following:
- Finance Insurance And Real Estate
- Services
- Retail Trade
- Manufacturing
- Transportation Communications Electric Gas And Sanitary Service
- Wholesale Trade
- Construction
- Mining
- Agriculture Forestry And Fishing

FIGIs

FIGIs are the emerging de facto standard for identifying securities in the financial industry as they are open-source and free to use. algoseek uses the Primary Exchange Composite FIGI when available. Otherwise, FIGI is used.

Multiple FIGIs

Some SecIds have multiple FIGI entries because the SecId represents a security that has changed over time. The FIGI column may have multiple entries representing the different entities. The FIGI values may be the same if a consistent Composite FIGI exists. In such a

case, the best practice is to use the last FIGI as your reference Id. The table on the next page shows examples of multiple FIGIs.

Multiple semicolons ";"

Some entries have a single FIGI with one or more ";" before or after the FIGI. This happens when there has been a change in the security but algoseek does not have a matching FIGI for it at a point when SecId is used. The best practice is to use the last FIGI listed for any security.

Table 4: SecIds with Multiple FIGIs

SecId	Tickers	FIGI
551883	QQQQ;QQQ	BBG000BSWKH7;BBG000BSWKH7;BBG000BSWKH7
1140088	QQQX	BBG000C0D2G9;BBG007SRSSN7
204368	BBG;HPR	BBG000BMYLC1;BBG00JPR7Y81

LOOKUP TABLE FILES

To find the correct SecId for a specific ticker, you will also need a trading date because sometimes the same ticker may refer to different SecIds due to the ticker being used by different companies during different periods (for example, S for Sprint Nextel and then Sprint). Also, sometimes to track the whole history for a specific security you will need to link two or more SecIds (for example, GOOG and GOOGL). It happens when the security structure changes. For that reason, algoseek invented a new unique identifier ASID that remains unchanged during the whole life of the security.

algoseek provides three versions of lookup files:

- **ticker_to_secid_lookup.csv** for mapping ticker name to SecId
- **figi_to_secid_lookup.csv** for FIGI to SecId conversion
- **equity_asid_lookup.csv** for ASID to SecId/SecIds mapping

Ticker to SecId Lookup

Table 5: Sample Data from Ticker to SecId Lookup File

Ticker	SecId	StartDate	EndDate
AABA	44754	20170619	20191004
AAC	32733	20070103	20100205

AAC	2276138	20101213	20121015
AAC	4307254	20141002	20191025
AACC	124678	20070103	20130613
AACG	763734	20191017	29991231

Table 5 above is a sample of a truncated Ticker-to-SecId lookup file, where AAC refers to three different securities during 20070103 - 20100205, 20101213 - 20121015 and 20141002 - 20191025.

Table 6 below summarizes the name, brief description, and data type for each data field (column) in the Ticker-to-SecId Lookup data file.

Table 6: Ticker to SecId Lookup File Fields Schema

Field	Format	Description
Ticker	string	Ticker symbol
SecId	integer	Unique ID per security
StartDate	string (yyyymmdd)	Start date for Ticker with this SecId
EndDate	string (yyyymmdd)	End date of ticker with this SecId. EndDate = 20991231 when the ticker is still being used

FIGI to SecId Lookup

Table 7: Sample Data from FIGI to Secid Lookup File

FIGI	SecId	StartDate	EndDate
BBG000002R06	481008	20070117	20100707
BBG000002R79	1397340	20081223	20131230
BBG000002RJ6	1184038	20070201	20120201
BBG000002RL3	1543109	20090716	20120607
BBG000002RQ8	488071	20070125	20130212

Table 7 is a sample of the FIGI-to-SecId lookup file, and Table 8 summarizes the name, brief description, and data type for each data field (column) in the FIGI-to-SecId Lookup file.

Table 8: FIGI to SecId Lookup File Fields Schema

Field	Format	Description
FIGI	string	FIGI identifier
SecId	integer	Unique ID per security
StartDate	string (yyyymmdd)	Start date for FIGI with this SecId
EndDate	string (yyyymmdd)	End date of FIGI with this SecId. EndDate = 20991231 when the FIGI is still being used

ASID to SecId Lookup

Table 9 demonstrates the full list of data fields in Equity ASID Lookup with sample contents for a few ASIDs, in which rows and columns are inverted for the convenience of the document display. For instance, one can see that ASID 1010000000001003 covers two SecIds (one SecId change event) for the same security during 2007-2022.

Table 9: Sample Data from Equity ASID Lookup

ASID	1010000000001000	1010000000001003	1010000000001006
LastTicker	A	AAA	AABA
Tickers	A	AAA	YHOO;AABA
TickersStartToEndDate	20070103:29991231	20200909:29991231	20070103:20170616; 20170619:20191004
SecIds	32952	6561562;7680133	44754
SecIdsStartToEndDate	20070103:29991231	20200909:20221014; 20221017:29991231	20070103:20191004

Table 10 below summarizes the name, brief description, and data type for each data field (column) in the Equity ASID Lookup file. The table column "Missing" indicates a default value or behavior in case the data field value is not present or cannot be determined where "Never" means that a value is always present in the data field.

Table 10: CSV File Fields Schema for Equity ASID Lookup

Field	Type (Format)	Missing	Description
ASID	string	Never	algoseek unique security identifier
LastTicker	string	Never	Current symbol name or the last available
Tickers	string (tricker1; ticker2;…)	Never	List of symbol names used. If security had its ticker changed, the field will have

			multiple tickers separated by a semicolon ";"
TickersStartToEndDate	string (yyyymmdd: yyyymmdd;…)	Never	Start and end dates for each ticker. EndDate = 20991231 when the ticker is still being used
SecIds	String (secid1;secid2;...)	Never	List of unique security identifiers (SecIds) related to security structure. If security had its SecId changed, the field will have multiple SecIds separated by a semicolon ";"
SecIdsStartToEndDate	string (yyyymmdd: yyyymmdd;…)	Never	Start and end dates for each SecId. EndDate = 20991231 when the SecId is still being used